PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 :
C12N 15/12, C07K 14/705, C12Q 1/68, G01N 33/68

A1

(11) International Publication Number: **WO 00/58461**

(43) International Publication Date: 5 October 2000 (05.10.00)

(21) International Application Number: PCT/EP00/02600

(22) International Filing Date: 23 March 2000 (23.03.00)

(30) Priority Data:
99106343.9 26 March 1999 (26.03.99) EP

(71)(72) Applicant and Inventor: RAPPOLD–HOERBRAND, Gudrun [DE/DE]; Hausackerweg 14, D–69118 Heidelberg (DE).

(74) Agent: KOESTER, Reinhold; Werderplatz 9, D–69120 Heidelberg (DE).

(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE; DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: GENE FOR ATAXIA

(57) Abstract

The present invention relates to the isolation, identification and characterization of a newly identified human gene responsible for disorders relating to ataxia, as well as the diagnosis and therapy of such disorders.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

GENE FOR ATAXIA

The present invention relates to the isolation, identification and characterization of newly identified human gene responsible for disorders relating to ataxia, as well as the diagnosis and therapy of such disorders.

Ataxia in general describes the inability of a patient to properly coordinate his or her movements and is caused by neuronal defects. The autosomal dominant forms mostly represent neurodegenerative disorders and are characterized by the selective loss of neurons within the brainstem and spinocerebellar tracts. Recessive forms of ataxia are often based on defects in neuronal cell communication, such as defective ion channels (Doyle and Stubbs, 1998).

Up to now 8 types of dominantly inherited ataxia have been genetically defined: Spinocerebellar Ataxia type 1-7 (SCA1-7) and Dentato-Rubral Pallido-Luysian Atrophy (DRPLA). Recent years have seen an impressive genetic contribution. The first gene for hereditary ataxia, SCAl, was discovered in 1993, and the gene of SCA7 was identified in 1996. Genetic mutations of SCA1, 2, 3, 6 and 7, have now been identified. For SCA4 and SCA5 the responsible genes have not been identified yet. For the SCA's with a known genetic mutation, individual genetic diagnosis and presymptomatic testing is possible.

Thus far, all known mutations for the different SCA's and for DRPLA consist of an unstable expansion of a CAG repeat in one of the gene's exons (reading frames). ‚CAG' stands for cytosine, adenine and guanine, the names of separate DNA molecules, which form a ‚triplet'. Dominant hereditary ataxia's are not the only disorders caused by an expanded triplet repeat. Among the disorders caused by an expanded triplet repeat are Huntington's disease, Friedreich's ataxia and Myotonic dystrophy.

In the SCA's, DRPLA, and in some other inherited diseases like Huntington's disease, a normally present CAG triplet repeat is expanded and has become instable. Translated into a protein, this expanded CAG triplet repeat results in a expanded tract of identical glutamine amino acids (a polyglutamine tract). Once the gene has been identified, the gene-products, proteins, can be studied. With one exception, SCA6, the proteins are previously unknown, novel proteins, with unknown functions. For SCA, the novel proteins are named ‚ataxin' with the corresponding SCA number, e.g. ataxin-l. Similarly

for DRPLA the protein is called 'atrophin', for Huntington's disease the protein is called ‚huntingtin' and for Friedreich´s ataxia, the protein is called ‚frataxin'.

SCA6 is the only Spinocerebellar Ataxia in which the function of the gene and protein is known. The protein in SCA6 is part of a voltage gated calcium channel in the cellular membrane, which plays a role in the cellular excitation. Other mutations in the gene for SCA6 can cause two other (allelic) disorders: hemiplegic migraine and episodic ataxia type II.

Combining the clinical information in the different Spinocerebellar Ataxia´s, some correlations become apparent. Within groups of SCA1, SCA2, SCA3, and SCA7 patients, the length of the CAG repeat shows a strong inverse correlation with the age at onset of symptoms, and also with progression of the disease. So, people with a later onset of symptoms usually have a smaller expansion of the CAG repeat. However, at any repeat length there is a rather wide variation in age at onset, of some twenty years or so. Thus, in a single individual the repeat length does not enable prediction of the age at onset or progression. Although essentially similar during an individual's lifetime, the CAG repeat is not stable during parent to child transmission, and may thus be different in parents and children. Especially in paternal transmissions, the expansion often increases. This in part explains anticipation, the phenomenon that the mean age at onset is earlier in next generations. In maternal transmission the repeat expansion is usually more stable.

Clinical manifestations vary considerably within each genetically defined type, and with exception of SCA7 and SCA6, only few global characteristics may be discerned for the SCA's, precluding an individual clinical diagnosis.

Research on histological and pathological aspects has focused mainly on the mechanism which causes cell death and ataxia. First, the presence of ataxin has been studied. As SCA's are dominantly inherited disorders, both a normal and a mutated SCA gene are present, and so both normal and mutated ataxin are expressed. It must be realised that genes are not active all the time, and that genes are turned on or off by signals according to cellular activity and specialisation. In the case of SCA1 and SCA2 for instance, it appears that ataxin-1 and ataxin-2 are present in many tissues and cells, and that the levels of ataxin do not correlate well with the presence of localised neurological cell death.

To explain the discrepancy between the presence of (mutated) ataxin and cell degeneration, interaction with other, possibly cell specific proteins has been suggested, and indeed, in Huntington's disease and various SCA's, aggregation with proteins has been demonstrated. Also, a tendency has been demonstrated for mutated ataxins to (self)aggregate, and proteins containing an expanded polyglutamine stretch have been found te be toxic for cultured cells.

Last year it was reported that in Huntington's disease and in SCA3 or Machado-Joseph Disease, certain brain cells showed inclusion bodies within the cell's nucleus. These inclusion bodies contained partially degraded molecules of mutated huntingtin and ataxin. Normally, huntingtin and ataxin are present in the cytoplasm of these cells, that is the fluid surrounding the nucleus, hut not within the nucleus. The presence of intranuclear inclusion bodies seemed to correlate with degeneration of these neurons. As these inclusion bodies also contained a protein degrading molecule, named ubiquitin, these inclusion bodies seem to indicate that the protein aggregate which they contain is resistant to degradation.

Friedreich's ataxia is the most frequent recessive ataxia occurring among white populations. But there are other recessive ataxias as well. They inelude: a) Ataxia-Telangiectasia (AT). This is an early onset ataxia with dilated veinules and telangiectasia. AT as it is known is almost as frequent as Freidreich's ataxia. b) Familial Vitamin E Deficiency, of which there are two forms, with or without fat malabsorption. The form without fat malabsorption is more common in the North African population, but rare everywhere else. Vitamin E supplementation is an effective treatment. Further known ataxias are c) Refsum Disease (ataxia with neuropathy and retinitis). d) Infantile Onset Spinocerebellar Ataxia, or IOSCA, a Finnish disease unknown in other world populations. e) Spastic Ataxia of Charlevoix-Saguenay. This is found mostly in Quebec, Canada. There exist further rare ataxias that don't have a name because they are not well charactenised. In two copies of the ataxia defect or mutation, one received from each parent. The parents of the patient are not affected because they carry only one copy. For all recessive diseases, the frequency of healthy carriers is much higher than the frequeney of patients, because the risk that two carriers of the same ataxia mutation marry is relatively low. For example, for Friedreich's ataxia, the frequency of healthy carriers is 1 in 120 while the frequeney of patients is 1 in 40,000. Almost all Friedreich's ataxia carriers and patients have the same ancestor (who lived most likely more than 10,000 years ago). This explains why Friedreich ataxia is more prevalent in

white populations and is almost non-existent among Japanese and black African populations.

The genetic defect in the gene leads to a lower expression of the protein called frataxin which is expressed about 10 to 20 times less compared to the expression in healthy persons. The important point is that there is still a little bit of frataxin made. This distinguishes the expansion mutation from other mutations, referred to as truncating mutations and frequently found in other recessive diseases which completely alter the production of the protein. Frataxin is a protein of the mitochondria, the structures of a cell where its fuel is produced from food. This energy conversion requires electron transport with iron containing proteins. Frataxin is present in tissues that are rich in mitochondria, which include the tissues that are affected in patients (dorsal root ganglia of the spinal cord, heart). Frataxin is present in every living organism, including fungi and yeast, because they also contain mitochondria. In yeast complete absence of frataxin causes iron accumulation within the mitochondria. In Friedreich's ataxia patients, a similar defect is also likely to apply, but to a lesser extent since there is still some residual frataxin present. Evidence for this are iron deposits and defect of some iron containing proteins in heart of patients. Mitochondrial iron accumulation seems therefore to be the culprit, since excess of iron is well known to stimulate the production of toxic compounds called free radicals on reactive oxygen species.

The present invention relates to the isolation and characterization of genes involved in constitution and regulation of the vertebrate nervous system, which is a major step towards elucidating the complex interactions required for an intact nervous system. The gene was identified via a positional cloning approach in the course of which the cytogenetic breakpoints of a patient suffering from severe cerebellar ataxia had been studied.

In order to investigate the breakpoint on the long arm of the X-chromosome, several YAC, PAC and cosmid clones (Vetrie et al., 1994) were chosen covering the region of CEPHy904G09223. DNA of these clones was used for FisH-analysis on metaphase spreads of the patient. It could be demonstrated that the PAC clone 1055C14 covers the breakpoint. The cosmids in close vicinity to the breakpoint-spanning PAC had already been sequenced by the Sanger Center in Cambridge.

Sequence analysis of Cosmid 9N5 allowed the design of appropriate primers for certain interesting regions. This allowed the amplifcation of parts of the potential new gene using a fetal brain cDNA library. The first seven exons of the gene could be identified. In order to obtain a full length cDNA clone, RACE experiments were performed and the resulting fragments were subcloned into a TOPO-vector (Invitrogen) and sequenced.

The complete cDNA has a coding region of 1251 bp encoding for 417 amino acids and a 3' untranslated region of 348 bp. The gene consists of nine exons ranging in size from 68 bp to 224 bp and is orientated from telomere (5'end) to centromere (3' end) (fig. 3). Exon 8 and 9 as well as the 3'UTR represent novel sequence data. The sequence as well as the translation into the potential protein are given in SEQ ID No. 1 and SEQ ID No. 2.

With respect to patient T.G., it was important to show whether the chromosomal breakpoint had affected this new gene. For this purpose all genomic fragments containing exons of the ataxia gene were isolated from PAC 1055C14, pooled and used for FisH-analysis on metaphase spreads of the patient's mother. The normal X-chromosome showed a signal in Xq22 plus a crosshybridization to Xq28. The rearranged X-chromosome showed a strong signal in Xp22, the crosshybridization signal and a weak signal in Xq22. This shows clearly that the breakpoint of the patient resides within the genomic locus of the ataxia gene.

To further verify this observation, Southern blot analysis was performed, using as a probe the entire ataxia cDNA. When using EcoRV as restriction enzyme, a band shift was observed in the patient, but not in healthy controls. This suggests, that indeed this novel gene is affected by the chromosomal breakage that has occured in the patient and might therefore contribute to the phenotype seen in the patient.

However, the gene responsible for the correct mediation of neuronal signalling can be used with respect to the diagnostic determination of other neurological diseases with yet unknown etiology, such as spastic paraplegia mutations in PLP, Pelizaeus-Merzbacher disease, diseases related to the vertebrate central nervous system and other various neurological diseases of yet unkown etiology.

The isolated genomic DNA or fragments thereof can be used for pharmaceutical purposes or as diagnostic tools or reagents for identification or characterization of the genetic defect involved in the disorders and diseases mentioned above.

Subject of the present invention are further ataxia proteins which are expressed after transcription of the ataxia gene into RNA or mRNA and which can be used in the therapeutic treatment of disorders related to mutations in said genes. The invention further relates to appropriate cDNA sequences which can be used for the preparation of recombinant proteins suitable for the treatment of such disorders.

Subject of the invention are further plasmid vectors for the expression of the DNA of these genes and appropriate cells containing such DNAs. It is a further subject of the present invention to provide means and methods for the genetic treatment of such disorders in the area of molecular medicine using an expression plasmid prepared by incorporating the nucleic acid molecules of this invention downstream from an expression promotor which effects expression in a mammalian host cell.

Brief Description of the SEQ ID:

SEQ ID NO. 1: Sequence of ataxia cDNA. Exon/Intron junctions are at the following positions: 72/73 (exon 1; size: 72 bp); 202/203 (exon 2; size: 130 bp); 270/271 (exon 3; size: 68 bp); 494/495 (exon 4; size: 224 bp); 575/576 (exon 5; size: 81 bp); 718/719 (exon 6; size: 143 bp); 933/934 (exon 7; size: 215 bp); 1083/1084 (exon 8; size: 150 bp); 1252 ATG stop codon (exon 9; size: 169 bp).

SEQ ID NO. 2: ataxia protein

SEQ ID NO. 3: exon 8 with neighbouring genomic sequences of the ataxia gene. Direction is from 5′ towards the 3′ end.

SEQ ID NO. 4: amino acid sequence for exon 8

SEQ ID NO. 5: exon 9 with neighbouring genomic sequences of the ataxia gene. Direction is from 5′ towards 3′ end.

SEQ ID NO. 6: part of the genomic sequence as published by the Sanger Center comprising exons 1 – 7. Exon/intron boundaries are at the following position of the sequence: exon1: 29850 – 29921 (72 bp); exon 2: 33026 – 33155 (130 bp); exon 3: 33445 – 33514 (68 bp); exon 4: 33752 – 33975 (224 bp); exon 5: 34115 – 34195 (81 bp); exon 6: 35760 – 35901 (143 bp); exon 7: 38782 – 38996 (215 bp). With respect to the published complementary strand, the exon/intron boundaries

(5'-boundary to 3'-boundary) are at the following positions of the sequence: exon 1: 9946 – 9875 (72 bp); exon 2: 6770 – 6641 (130 bp); exon 3: 6351 – 6282; (68 bp); exon 4: 6044 – 5821 (224 bp); exon 5: 5681 – 5601 (81 bp); exon 6: 4036 – 3895 (143 bp); exon 7: 1014 – 800 (215 bp).

Brief Description of the figures:

Fig. 1: Schematic presentation of the genomic organization of the gene. Exons 1 - 7 were already sequenced by genome-wide sequencing efforts. Exons 8 and 9 as well as the 3'UTR have been sequenced according to the present invention.

Since the target gene leading to ataxia was unknown prior to the present invention, the biological and clinical association of patients with this deletion could give insights to the function of this gene. In the present study, fluorescence in situ hybridization (FISH) was used to examine metaphase and interphase lymphocyte nuclei of patients.

Subject of the present invention are therefore DNA sequences or fragments thereof which are part of the genes responsible for ataxia. DNA sequences or fragments of this gene, as well as the respective full length DNA sequence of this gene can be transformed in an appropriate vector and transfected into cells. When such vectors are introduced into cells in an appropriate way as they are present in healthy humans, it is devisable to treat diseases of patients suffering from various types of ataxia by modern means of gene therapy. For example, ataxia can be treated by removing the respective mutated gene. It is also possible to stimulate the respective genes which compensate the action of the genes responsible for ataxia, i.e. by inserting DNA sequences before, after or within the ataxia gene in order to increase the expression of the healthy allels. By such modifications of the genes, the ataxia gene become activated or silent, respectively. This can be accomplished by inserting DNA sequences at appropriate sites within or adjacent to the gene, so that these inserted DNA sequences interfere with the ataxia gene and thereby activate or prevent their transcription. It is also devisable to insert a regulatory element (e.g. a promotor sequence) before said ataxia gene to stimulate the gene to become active. It is further devisable to stimulate the respective promotor sequence in order to overexpress the healthy functional allele. The modification of genes can be generally achieved by inserting exogenous DNA sequences into the ataxia gene via homologous recombination.

The DNA sequences according to the present invention can also be used for transformation of said sequences into animals, such as mammals, via an appropriate vector system. These transgenic animals can then be used for *in vivo* investigations for screening or identifying pharamceutical agents which are useful in the treatment of diseases involved with defects of the ataxia gene. If the animals positively respond to the administration of a candidate compound or agent, such agent or compound or derivatives thereof would be devisable as pharmaceutical agents. By appropriate means, the DNA sequences of the present invention can also be used in genetic experiments aiming at finding methods in order to compensate for the loss of genes responsible for ataxia (knock-out animals).

In a further object of this invention, the DNA sequences can also be used to be transformed into cells. These cells can be used for identifying pharmaceutical agents useful for the treatment of diseases involved with ataxia, or for screening of such compounds or library of compounds, especially using high-throuput screening systems. In an appropriate test system, variations in the phenotype or in the expression pattern of these cells can be determined, thereby allowing the identification of interesting candidate agents in the development of pharmaceutical drugs.

The DNA sequences of the present invention can also be used for the design of appropriate primers which hybridize with segments of the ataxia gene or fragments thereof under stringent conditions. Appropriate primer sequences can be constructed which are useful in the diagnosis of people who have a genetic defect causing ataxia. Such primers comprise nucleic acid sequences having a length of 10 – 40 nucleotides, preferably 15 – 30 nucleotides. Especially preferred primers are given in example 3.

In general, DNA sequences according to the present invention are understood to embrace also such DNA sequences which have at least 80 %, preferably at least 90 %, or at least 95 % or 98 % sequence identity to the specific DNA sequences according to the present invention. This also includes such nucleic acid sequences which are degenerate to the specific sequences shown, based on the degeneracy of the genetic code, or fragments thereof which hybridize under stringent conditions with the specifically shown DNA sequences. Hybridisation is to be understood to take place at high stringency conditions using Church buffer (0.5 M NaPi pH 7.2, 7% SDS, 1 mM EDTA) at 65°C and by washing in 40 mM NaPi, 1% SDS at 65°C.

In principle, all oligonucleotide primers and probes for amplifying and detecting a genetic defect responsible for ataxia in a biological sample are suitable for amplifying a target ataxia associated sequence. Especially, suitable exon specific primer pairs can be provided. Subsequently, a suitable detection, e.g. a radioactive or non-radioactive label is carried out.

Also, a single stranded RNA can be used as target. Methods for reversed transcribing RNA into cDNA are also well known and described in Sambrook et al., Molecular Cloning: A Laboratory Manual, New York, Cold Spring Harbor Laboratory 1989. Alternatively, preferred methods for reversed transcription utilize thermostable DNA polymerases having RT activity.

Further, the technique described before can be used for selecting those person from a group of persons being of short stature characterized by a genetic defect and which allows as a consequence a more specific medical treatment.

In another subject of the present invention, the ataxia protein can be used as pharmaceutical agents. These proteins initiate a still unknown cascade of biological effects on a molecular level involved with ataxia. They can be used in the treatment of ataxia, and related diseases based on the same etiology. Within the meaning of the present invention, the term "ataxia protein" also comprises proteins and functional fragments thereof, which have a similar or comparable physiological effect as the ataxia protein described in SEQ ID. No. 2. Such proteins comprise modifications, deletions, substitutions or variations of one or more amino acids, whereby the resulting variant has a homology of at least 80 %, preferably at least 90 %, 95 % or 99 % of the protein described in SEQ ID. 2. The ataxia proteins have a length of 300 – 600 amino acids, preferably 350 – 500 amino acids, and most preferably of 400 – 450 amino acids.

As used herein, the term „isolated" refers to the original derivation of the DNA molecule by cloning. It is to be understood however, that this term is not intended to be so limiting and, in fact, the present invention relates to both naturally occurring and synthetically prepared seqences, as will be understood by the skilled person in the art.

The DNA molecules of this invention may be used in forms of gene therapy involving the use of an expression plasmid prepared by incorporating an appropriate DNA sequence of this invention downstream from an expression promotor that effects

expression in a mammalian host cell. Suitable host cells are procaryotic or eucaryotic cells. Procaryotic host cells are, for example, E. coli, Bacillus subtilis, and the like. By transfecting host cells with replicons originating from species adaptable to the host, that is, plasmid vectors containing replication starting point and regulator sequences, these host cells can be transfected with the desired gene or cDNA. Such vectors are preferably those having a sequence that provides the transfected cells with a property (phenotype) by which they can be selected. For example, for E. coli hosts the strain E. coli K12 is typically used, and for the vector either pBR322 or pUC plasmids can be generally employed. Examples for suitable promotors for E. coli hosts are trp promotor, lac promotor or lpp promotor. If desired, secretion of the expression product through the cell membrane can be effected by connecting a DNA sequence coding for a signal peptide sequence at the 5′ upstream side of the gene. Eucaryotic host cells include cells derived from vertebrates or yeast etc.. As a vertebrate host cell, COS cells can be used (Cell, 1981; 23: 175 - 182), or CHO cells. Preferably, promotors can be used which are positioned 5′ upstream of the gene to be expressed and having RNA splicing positions, polyadenylation and transcription termination seqences.

The present invention is illustrated by the following examples.

Example 1
Patient

Patient T.G. is a ten year old boy suffering from slight mental retardation and severe cerebellar ataxia. In a MRI-scan severe hypo- and dysmyelinisation was observed infra- and supra- tentorial. Cytogenetically, the boy, as well as his healthy mother, exhibit an inversion of the X-chromosome with the breakpoints residing in Xp22 and Xq22, respectively.

Example 2
Identification of the ataxia gene

a) Florescence in situ hybridization (FISH)

FISH studies using appropriate cosmids were carried out according to published methods (Lichter and Cremer, 1992). In short, one microgram of the respective cosmid

clone was labeled with biotin and hybridized to human metaphase chromosomes under conditions that suppress signals from repetitive DNA sequences. Detection of the hybridization signal was via FITC-conjugated avidin. Images of FITC were taken by using a cooled charge coupled device camera system (Photometrics, Tucson, AZ).

b) Physical mapping

In order to investigate the breakpoint on the long arm of the X-chromosome, several YAC, PAC and cosmid clones (Vetrie et al., 1994) were chosen covering the region of CEPHy904G09223. DNA of these clones was used for FisH-analysis on metaphase spreads of the patient. It could be demonstrated that the PAC clone 1055C14 covers the breakpoint. The cosmids in close vicinity to the breakpoint-spanning PAC had already been sequenced by the Sanger Center in Cambridge.

c) Southern Blot Hybridisation

Southern blot hybridisations were carried out at high stringency conditions in Church buffer (0.5 M NaPi pH 7.2, 7% SDS, 1mM EDTA) at 65°C and washed in 40 mM NaPi, 1% SDS at 65°C.

d) FISH Analysis

Biotinylated cosmid DNA (insert size 32 - 45 kb) or cosmid fragments (10 - 16 kb) were hybridised to metaphase chromosomes from stimulated lymphocytes of patients under conditions as described previously (Lichter and Cremer, 1992). The hybridised probe was detected via avidin-conjugated FITC.

e) PCR Amplification

All PCRs were performed in 50 µl volumes containing 100 pg-200 ng template, 20 pmol of each primer, 200 µM dNTP's (Pharmacia), 1.5 mM $MgCl_2$, 75 mMTris/HCl pH9, 20mM $(NH_4)_2SO_4$ 0.01% (w/v) Tween20 and 2 U of Goldstar DNA Polymerase (Eurogentec). Thermal cycling was carried out in a Thermocycler GeneE (Techne).

f) Exon Amplification

For exon amplification experiments, the methods as previously described in Church et al., 1994 were used.

Example 3

Primer for the amplification of ataxia exons

Exons 2 and 3 as well as exons 4 and 5 were amplified in one PCR-reaction, as the introns are rather small.

GLRA4 Exon 1 For : 5'-ATG ACA ACT CTT GTT CCT GCA ACC CTC TCC-3'
GLRA4 Exon 1 Rev : 5'-TGG CTA GTG TTT GCA TGC ACC-3'
PCR conditions : 30'' 94°C, 30'' 58°C, 30'' 72°C, 38 cycles

GLRA4 Exon 2 For : 5'-GCA CAT AAC TGG CCT CAG ACT T-3'

GLRA4 Exon 3 Rev : 5'-CAC TCA ACA TAG GCT GGA AGT-3'
PCR conditions : 30'' 94°C, 30'' 60°C, 45'' 72°C, 38 cycles

GLRA4 Exon 4 For : 5'-CCT GAG ATG TGT TCC CAA CAT-3'

GLRA4 Exon 5 For : 5'-CCA GTA AGC CGA TGT CAC TTC-3'
GLRA4 Exon 5 Rev : 5'-GTT ATT CCA GGC TCT CTG TGA-3'
PCR conditions : 30'' 94°C, 30'' 59°C, 30'' 72°C, 38 cycles

GLRA4 Exon 6 For : 5'-CAG CAT CCA TAC TCT GCA GC-3'
GLRA4 Exon 6 Rev : 5'-AGG TTC TCC TGT GGC TCA CA-3'
PCR conditions : 30'' 94°C, 30'' 60°C, 30'' 72°C, 38 cycles

GLRA4 Exon 7 For : 5'-TCA GGC TCA GCT ACA GGC TG-3'
GLRA4 Exon 7 Rev : 5'-GGT ACT CTA TGG CAA GCA AGT T-3'
PCR conditions : 30'' 94°C, 30'' 58°C, 30'' 72°C, 38 cycles

GLRA4 Exon 8 For : 5'-GTG TCC TAC GTG AAG GCA AT-3'
GLRA4 Exon 8 Rev : 5'-TCC AAG CGT TGG CGC CTC T-3'
PCR conditions : 30'' 94°C, 30'' 58°C, 30'' 72°C, 38 cycles

GLRA4 Exon 9 For : 5'-CCA AGG CAC CTT GTC TGC ATA ACA-3'
GLRA4 Exon 9 Rev : 5'-GGA GAT GGT GTC AAT TCT CTT GGC-3'
PCR conditions : 30'' 94°C, 30'' 59°C, 30'' 72°C, 38 cycles

Claims

1. An isolated human nucleic acid sequence encoding a human ataxia protein having essentially the amino acid sequence of SEQ ID NO: 2.

2. A nucleic acid sequence according to claim 1 having the cDNA sequence according to SEQ ID NO: 1 or sequences substantially complementary to said cDNA sequence or fragments that are substantially complementary to fragments of said cDNA sequence.

3. A nucleic acid sequence according to claim 1 which is a genomic sequence.

4. A nucleic acid sequence according to claims 1 – 3 comprising the SEQ ID NO: 3.

5. A nucleic acid sequence according to claims 1 – 4 comprising the SEQ ID NO: 5.

6. A nucleic acid sequence according to claim 3 and comprising the SEQ ID NO: 6.

7. A recombinant host cell containing a nucleic acid sequence according to any of claims 1 – 6.

8. A vector comprising a nucleic acid sequence according to any of claims 1 – 6.

9. A protein encoded by any of the sequences according to claim 1 – 6 or a functional fragment or variant thereof.

10. A protein having the amino acid sequence selected from the group consisting of a) the amino acid sequence shown in SEQ ID NO. 2; b) a portion of the amino acid sequence shown in SEQ ID NO. 2; c) amino acid sequences which are homologous variants of the amino acid sequence shown in SEQ ID NO. 2; or d) portions of amino acid sequences which are homologous variants of the amino acid sequence shown in SEQ ID NO. 2.

11. A recombinant nucleic acid sequence comprising a genomic or cDNA sequence that encodes a protein according to claim 10.

12. A recombinant nucleic acid sequence according to claim 11 wherein said genomic DNA or cDNA sequence is operatively linked to an expression control sequence in said nucleic acid molecule.

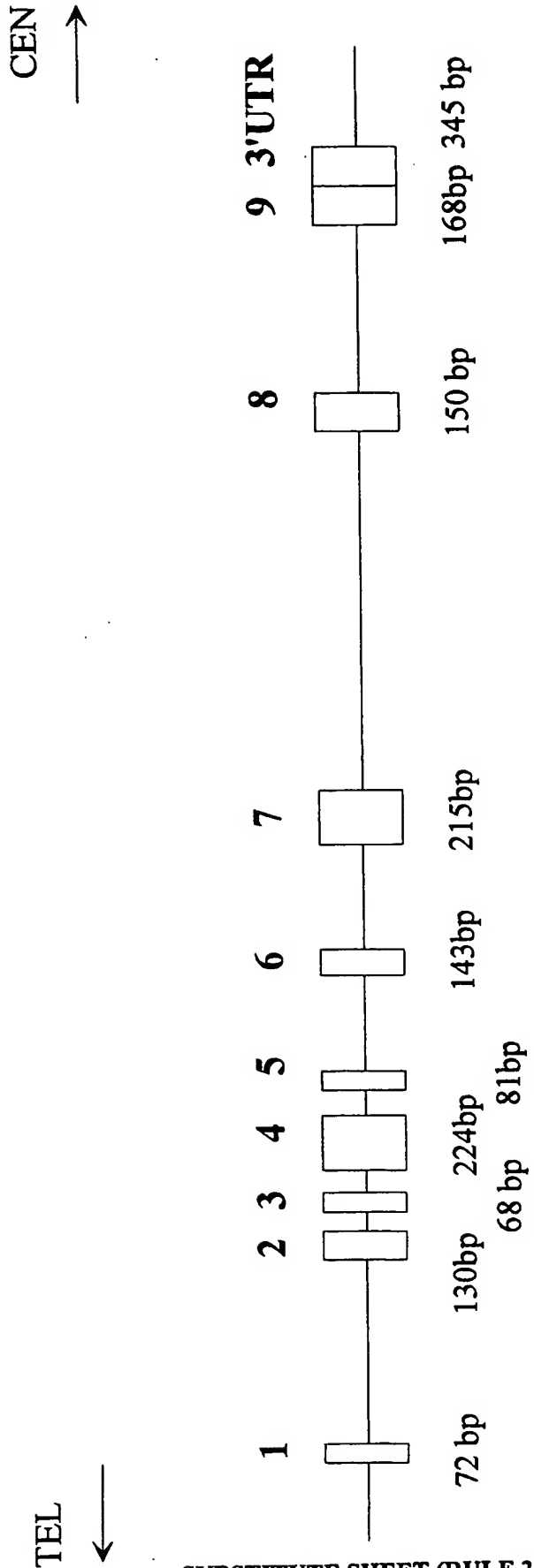
13. A method of detecting a genetic defect in the gene responsible for ataxia wherein the gene comprises an nucleic acid sequence according to claim 1, said method comprising the steps of a) amplifying one or more fragments of said gene and the ataxia gene known to be normal by the polymerase chain reaction (PCR) with the same or substantially the same PCR primers; b) determining whether said ataxia gene contains any mutations or genetic defects by comparing the PCR products of the amplification of said gene with those from the amplification of the normal gene, and c) detecting differences between the PCR products associated with mutations.

14. A method according to claim 13, wherein the gene comprises a nucleic acid sequence according to SEQ ID No. 3 (exon 8) or fragments thereof.

15. A method according to claim 13 or 14, wherein the gene comprises a nucleic acid sequence according to SEQ ID No. 5 (exon 9) or fragments thereof.

16. A method for identifying or screening of candidates for pharmaceutical agents useful for the treatment of disorders relating to mutations in the ataxia gene comprising providing a test system containing a transformed host cell according to claim 7 and determining variations in the phenotype of said cells or variations in the expression products of said cells after contacting said cells with said candidate pharmaceutical agents.

FIG. 1


CEN
TEL
1
2 3
4
5
6
7
8
9 3'UTR
72 bp
130bp
224bp
68 bp
81bp
143bp
215bp
150 bp
168bp 345 bp

SEQUENCE LISTING

<110> Rappold Dr., Gudrun

<120> Ataxia gene

<130> ataxia_ba

<140>
<141>

<160> 6

<170> PatentIn Ver. 2.0

<210> 1
<211> 1600
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: cDNA

<220>
<221> CDS
<222> (1)..(1251)

<400> 1

```
atg aca act ctt gtt cct gca acc ctc tcc ttc ctt ctt ctc tgg acc   48
Met Thr Thr Leu Val Pro Ala Thr Leu Ser Phe Leu Leu Leu Trp Thr
  1               5                  10                  15

ctg cca ggg cag gtc ctc ctc agg gtg gcc ttg gca aaa gag gaa gtc   96
Leu Pro Gly Gln Val Leu Leu Arg Val Ala Leu Ala Lys Glu Glu Val
             20                  25                  30

aaa tct gga acc aag ggg tcc cag ccc atg tcc ccc tct gat ttc cta  144
Lys Ser Gly Thr Lys Gly Ser Gln Pro Met Ser Pro Ser Asp Phe Leu
         35                  40                  45

gac aaa ctt atg ggg cga aca tct gga tat gat gcc agg att cgg ccc  192
Asp Lys Leu Met Gly Arg Thr Ser Gly Tyr Asp Ala Arg Ile Arg Pro
     50                  55                  60

aat ttt aaa ggc cca ccc gtg aac gtg acc tgc aac atc ttc atc aac  240
Asn Phe Lys Gly Pro Pro Val Asn Val Thr Cys Asn Ile Phe Ile Asn
 65                  70                  75                  80

agt ttc agc tcc atc acc aag acc aca atg gac tac cgg gtg aat gtc  288
Ser Phe Ser Ser Ile Thr Lys Thr Thr Met Asp Tyr Arg Val Asn Val
                 85                  90                  95

ttc ttg cgg caa cag tgg aat gac cca cgc ctg tcc tac cga gaa tat  336
Phe Leu Arg Gln Gln Trp Asn Asp Pro Arg Leu Ser Tyr Arg Glu Tyr
            100                 105                 110

cct gat gac tct ctg gac ctc gat ccc tcc atg ctg gac tct atc tgg  384
Pro Asp Asp Ser Leu Asp Leu Asp Pro Ser Met Leu Asp Ser Ile Trp
        115                 120                 125
```

```
aag cca gac ctc ttc ttt gct aat gag aaa ggg gcc aac ttc cat gag   432
Lys Pro Asp Leu Phe Phe Ala Asn Glu Lys Gly Ala Asn Phe His Glu
    130                 135                 140

gtg acc acg gac aac aag tta ctg cgc atc ttc aag aat ggg aat gtg   480
Val Thr Thr Asp Asn Lys Leu Leu Arg Ile Phe Lys Asn Gly Asn Val
145                 150                 155                 160

ctg tac agc atc agg ctg acc ctc att ttg tcc tgc ctg atg gac ctc   528
Leu Tyr Ser Ile Arg Leu Thr Leu Ile Leu Ser Cys Leu Met Asp Leu
                165                 170                 175

aag aac ttc ccc atg gac atc cag acc tgc acg atg cag ctt gag agg   576
Lys Asn Phe Pro Met Asp Ile Gln Thr Cys Thr Met Gln Leu Glu Arg
            180                 185                 190

ttt ggc tac acc atg aaa gac ctc gtg ttt gag tgg ctg gaa gat gct   624
Phe Gly Tyr Thr Met Lys Asp Leu Val Phe Glu Trp Leu Glu Asp Ala
        195                 200                 205

cct gct gtc caa gtg gct gag ggg ctg act ctg ccc cag ttt atc ttg   672
Pro Ala Val Gln Val Ala Glu Gly Leu Thr Leu Pro Gln Phe Ile Leu
    210                 215                 220

cgg gat gag aag gat cta ggc tgt tgt acc aag cac tac aac aca ggg   720
Arg Asp Glu Lys Asp Leu Gly Cys Cys Thr Lys His Tyr Asn Thr Gly
225                 230                 235                 240

aaa ttc acc tgc atc gag gta aag ttt cac ctg gaa cgg cag atg ggc   768
Lys Phe Thr Cys Ile Glu Val Lys Phe His Leu Glu Arg Gln Met Gly
                245                 250                 255

tac tat ctg att cag atg tac atc ccc agc cta ctc atc gtc atc ctg   816
Tyr Tyr Leu Ile Gln Met Tyr Ile Pro Ser Leu Leu Ile Val Ile Leu
            260                 265                 270

tcc tgg gtc tcc ttc tgg atc aac atg gat gct gcc cct gcc cgt gtg   864
Ser Trp Val Ser Phe Trp Ile Asn Met Asp Ala Ala Pro Ala Arg Val
        275                 280                 285

ggc ctg ggc atc acc acc gtg ctc acc atg acc acc cag agc tct ggc   912
Gly Leu Gly Ile Thr Thr Val Leu Thr Met Thr Thr Gln Ser Ser Gly
    290                 295                 300

tcc cgg gcc tct ttg cct aag gtg tcc tac gtg aag gca atc gac atc   960
Ser Arg Ala Ser Leu Pro Lys Val Ser Tyr Val Lys Ala Ile Asp Ile
305                 310                 315                 320

tgg atg gct gtg tgt ctg ctc ttt gtg ttc gct gcc ttg ctg gag tat  1008
Trp Met Ala Val Cys Leu Leu Phe Val Phe Ala Ala Leu Leu Glu Tyr
                325                 330                 335

gct gcc ata aat ttt gtt tct cgt cag cat aaa gaa ttc ata cga ctt  1056
Ala Ala Ile Asn Phe Val Ser Arg Gln His Lys Glu Phe Ile Arg Leu
            340                 345                 350

cga aga agg cag agg cgc caa cgc ttg gag gaa gat atc atc caa gaa  1104
Arg Arg Arg Gln Arg Arg Gln Arg Leu Glu Glu Asp Ile Ile Gln Glu
        355                 360                 365

agt cgt ttc tat ttc cgt ggc tat ggc ttg ggc cac tgc ctg cag gca  1152
```

Ser Arg Phe Tyr Phe Arg Gly Tyr Gly Leu Gly His Cys Leu Gln Ala
370 375 380

aga gat gga ggt cca atg gaa ggt tct ggc att tat agt ccc caa cct 1200
Arg Asp Gly Gly Pro Met Glu Gly Ser Gly Ile Tyr Ser Pro Gln Pro
385 390 395 400

cca gcc cct ctt cta agg gaa gga gaa acc acg cgg aaa ctc tac gtg 1248
Pro Ala Pro Leu Leu Arg Glu Gly Glu Thr Thr Arg Lys Leu Tyr Val
405 410 415

gac tgagccaaga gaattgacac catctcccaa gggcgaattc cctttcactt 1301
Asp

tcctcatctt caatatcttc tactgggttg tctataaagt gctacggtca gaagatatcc 1361

accaggctct gtgaataggg tgggagctat agagtcctgc tgctggcctc ctgcttcctc 1421

tgggtgggct ttctccctca gttagactcc attaggggtt tggacagttc cttcctgatc 1481

tcccactcag aacttcaact accagtccca aagctatgtg ggcctatatt gcatggtgcc 1541

aatggtggct gtacttataa agatggctta tctacccctа aaaaaaaaac aaaaaaaag 1600

<210> 2
<211> 417
<212> PRT
<213> Artificial Sequence

<400> 2
Met Thr Thr Leu Val Pro Ala Thr Leu Ser Phe Leu Leu Leu Trp Thr
1 5 10 15

Leu Pro Gly Gln Val Leu Leu Arg Val Ala Leu Ala Lys Glu Glu Val
20 25 30

Lys Ser Gly Thr Lys Gly Ser Gln Pro Met Ser Pro Ser Asp Phe Leu
35 40 45

Asp Lys Leu Met Gly Arg Thr Ser Gly Tyr Asp Ala Arg Ile Arg Pro
50 55 60

Asn Phe Lys Gly Pro Pro Val Asn Val Thr Cys Asn Ile Phe Ile Asn
65 70 75 80

Ser Phe Ser Ser Ile Thr Lys Thr Thr Met Asp Tyr Arg Val Asn Val
85 90 95

Phe Leu Arg Gln Gln Trp Asn Asp Pro Arg Leu Ser Tyr Arg Glu Tyr
100 105 110

Pro Asp Asp Ser Leu Asp Leu Asp Pro Ser Met Leu Asp Ser Ile Trp
115 120 125

Lys Pro Asp Leu Phe Phe Ala Asn Glu Lys Gly Ala Asn Phe His Glu
130 135 140

Val Thr Thr Asp Asn Lys Leu Leu Arg Ile Phe Lys Asn Gly Asn Val
145 150 155 160

Leu Tyr Ser Ile Arg Leu Thr Leu Ile Leu Ser Cys Leu Met Asp Leu
165 170 175

Lys Asn Phe Pro Met Asp Ile Gln Thr Cys Thr Met Gln Leu Glu Arg
180 185 190

Phe Gly Tyr Thr Met Lys Asp Leu Val Phe Glu Trp Leu Glu Asp Ala
195 200 205

Pro Ala Val Gln Val Ala Glu Gly Leu Thr Leu Pro Gln Phe Ile Leu
210 215 220

Arg Asp Glu Lys Asp Leu Gly Cys Cys Thr Lys His Tyr Asn Thr Gly
225 230 235 240

Lys Phe Thr Cys Ile Glu Val Lys Phe His Leu Glu Arg Gln Met Gly
245 250 255

Tyr Tyr Leu Ile Gln Met Tyr Ile Pro Ser Leu Leu Ile Val Ile Leu
260 265 270

Ser Trp Val Ser Phe Trp Ile Asn Met Asp Ala Ala Pro Ala Arg Val
275 280 285

Gly Leu Gly Ile Thr Thr Val Leu Thr Met Thr Thr Gln Ser Ser Gly
290 295 300

Ser Arg Ala Ser Leu Pro Lys Val Ser Tyr Val Lys Ala Ile Asp Ile
305 310 315 320

Trp Met Ala Val Cys Leu Leu Phe Val Phe Ala Ala Leu Leu Glu Tyr
325 330 335

Ala Ala Ile Asn Phe Val Ser Arg Gln His Lys Glu Phe Ile Arg Leu
340 345 350

Arg Arg Arg Gln Arg Arg Gln Arg Leu Glu Glu Asp Ile Ile Gln Glu
355 360 365

Ser Arg Phe Tyr Phe Arg Gly Tyr Gly Leu Gly His Cys Leu Gln Ala
370 375 380

Arg Asp Gly Gly Pro Met Glu Gly Ser Gly Ile Tyr Ser Pro Gln Pro
385 390 395 400

Pro Ala Pro Leu Leu Arg Glu Gly Glu Thr Thr Arg Lys Leu Tyr Val
405 410 415

Asp

<210> 3
<211> 441
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: exon 8 of ataxia gene with neighbouring genomic sequences

<220>
<221> CDS
<222> (173)..(322)

<400> 3
tccatctggg ctttcagaca atgggatatg tcatggaagg cttctttaaa caccagaaga 60

aattcaggat aaagctcaaa aagagcaggc aatcgatagg ggttgaaaat ccactcagta 120

ggccacggaa ggacttcaag aaggttgatc gttctgtcgc tggatgttgt ag gtg tcc 178
Val Ser
1

tac gtg aag gca atc gac atc tgg atg gct gtg tgt ctg ctc ttt gtg 226
Tyr Val Lys Ala Ile Asp Ile Trp Met Ala Val Cys Leu Leu Phe Val
5 10 15

ttc gct gcc ttg ctg gag tat gct gcc ata aat ttt gtt tct cgt cag 274
Phe Ala Ala Leu Leu Glu Tyr Ala Ala Ile Asn Phe Val Ser Arg Gln
20 25 30

cat aaa gaa ttc ata cga ctt cga aga agg cag agg cgc caa cgc ttg 322
His Lys Glu Phe Ile Arg Leu Arg Arg Arg Gln Arg Arg Gln Arg Leu
35 40 45 50

gtgaggtaca actagggagc ggcccactgc ctttctcctt tggactcctt cctccctagc 382

tccgcctccc cacaaagccg agacccaccc cagggctgca agggactaga atagggcaa 441

<210> 4
<211> 50
<212> PRT
<213> Artificial Sequence

<400> 4
Val Ser Tyr Val Lys Ala Ile Asp Ile Trp Met Ala Val Cys Leu Leu
1 5 10 15

Phe Val Phe Ala Ala Leu Leu Glu Tyr Ala Ala Ile Asn Phe Val Ser
20 25 30

Arg Gln His Lys Glu Phe Ile Arg Leu Arg Arg Arg Gln Arg Arg Gln
35 40 45

Arg Leu
50

<210> 5
<211> 534
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: exon 9 of
ataxia gene with neighbouring genomic sequences

<220>
<221> conflict
<222> (319)..(488)

<223> exon 9

<400> 5
gagatgtcct ctcacagtgg ccggcacttc actggatttt agataactct aaatgatgtt 60
ttgcatttta catgggtagc ttagagccta tctgccatgg aatcatgaca tacactatat 120
atagcacgta gagctgctta gcatacatct atctcatgag tagtgttgtg accataaaaa 180
gcatgcaatg ggcctgtgta cccaaggcac cttgtctgca taacattccc ctatgagata 240
ttgatgggtg gtctgttctc tatccctctg cctttgagct ttgacctcaa attctgccgt 300
tctcctgtaa tttcccagga ggaagatatc atccaagaaa gtcgtttcta tttccgtggc 360
tatggcttgg gccactgcct gcaggcaaga gatggaggcc aatggaaggt tctggcattt 420
atagtcccca acctccagcc cctcttctaa gggaaggaga aaccacgcgg aaactctacg 480
tggactgagc caagagaatt gacaccatct cccaagggcg aattcccttt cact 534

<210> 6
<211> 39796
<212> DNA
<213> human

<400> 6
gatctctttc actgtcataa ttttatcact gttaaaattt ttgcaaaggc agtttcaaag 60
aggtcctctc tactgaggac ctctttagaa gtctttgtgg gatgggtaat gtccctcttc 120
cagactcaag ccctcagctc caacctcctc tggagtcctc tgagctgacc gctgtggctg 180
agctgggcca ggggtatgct gcctccctcc actaccagtg tccagttgaa tgcacaagga 240
ccagatgtgt attcttcaag tgtgagccaa tctggtgcca ctgaatcctg caaagaggca 300
tgcccaggcc ctggaaatgt tatagtaggt agtcagacat gagcagggca ggagagggct 360
cccccaaaac taggaatgtc aggtgaccat caggtgatgg tcaggtgatt gttaaactgt 420
ctctctaaaa taataattgg ttgcagttgg caccagggaa tgtctcccaa tagatagaaa 480
aacctaaaac tggtgatcag cagcttcctg ataagatctc aggagctggg ctagtgggtt 540
caagcatgtg cactaagagg caaaatggca gagtttaact ggtatatggc cttcctctag 600
aacactcaac tgataaggaa agaacgcctc atgtgagcat gcttacaact tcagtaaacg 660
tactgtgcat acggcccctc ccaaatgctg gcaggccagt gcacatgtgg acagcccacc 720
ccaagggaag aatcaagaga gaagagatgc aaacccccag aagcgtgcca acatataaaa 780
ccccaagtcg aaggtcaaac catgcatttg aacctctcaa gtcgcctgct tggtcttctt 840
acaagtatac gttacatcct tttgttccta cactaaaact tcttcagaaa ctttcactcc 900
tcctctaaaa cttgcctcag tctttccctt tgccttatga ccctccatcg aattctttct 960
tctgaggagt cgagaattga ggttgctgca gacccgtaca gatttgccac tagcaggaag 1020

accagaagtt gcaacagctg caggcagaga gtaggggtga ggtgctcagt gcaaagagaa 1080

ttctggtggc aaggggatcc ttgatctgtg gtggtgtgtg accactctgg gcagtgatcc 1140

aaaaaattga attcacacac actcacttct tcctggagca agcagaacaa atgtttcttg 1200

catcctgctc agctccagcc ctcctcaagc tccaaccacc ccccaccaca ccaaataaaa 1260

aaaaaagaga tttaaagctc agctattttt aatttttgc ttctccaaaa gttggcctgg 1320

agctcacagc tatttttcct cttttctcta cactggagtt ttcatcgatt tgaaactctc 1380

ttaatccttt aggccatggg ctgatgaacc ttctgaggta acctctttgg caaactgcat 1440

aggtgggtgt cagtggggat catgctcttg ttgagggcaa gaagggcctt cctcagccct 1500

tgcaagacaa gactccctct ctgacccttt tctggcattt cctgctgttc ttggctatac 1560

tattcctaaa atctttacaa tcttcaaagt gatgcttttt tgtatgctaa tgagttgact 1620

tatggctggc agcaggtagg tagcctcagg atgggagttg atcaccaaaa agaccaaggc 1680

agaattagag gattgtgatg ttcagccaca cccccagcct cctaggaggg gagaagggct 1740

gaaagttaag ttaatcactg atggccaatg atttaatcaa tcaagcctac ataataaagc 1800

cttcataaaa acccagaaag acagggtttg aaaagcctgt ggatagccat acatggagtt 1860

tgaacatctg cccaccagga ggatggcata tcccaactcc acagagacag aagctcctgt 1920

gctcaggact ttagctctca ccctgtgtat ctcttcatct ggccattcat ttgtatcctt 1980

taaaatctcc tttctaactg gtaaaattaa gtaagtgctt ccctgagttc tgtgagctgc 2040

tctagcaaac tgatcaaacc tgcggaaggg ttcatgggaa ccccgattta tagccagtca 2100

gtcagaagca gaggcaaaat aacctggggc ttgtgattgg cactggaagt gggagacagg 2160

cttggggact gagccctcaa cctgcaggat ctgacgctat ctccaggtag gtagcatcag 2220

aattgaatcg gaggacatcc agctggtgtc tcctgcagag ttgcttgctt ggtgtgtgag 2280

aaattgggaa acacaccaca catttggtca cagaagtgtt ccatgttgaa tgttgttggg 2340

taagagcata agaataacag cttatgtttt ttttctactc aggcatctac ctcttcccca 2400

aatctagttc cttctgccct ctgtgtggag aagactactg ggccttaagc tcctgaagga 2460

atcacagcat tggagccagc acaagcagca cttcctagaa atctccaagg cttcagagct 2520

gatcacactt ggctagggct gagcccaagt tccatggaaa gaaaccttta caaggtgcca 2580

tggcacctgc cttaaagagg gcaggcaccc tcatcagggc aggcaactaa taggagtccc 2640

tggcagcttg ggagttggaa acaaatctta gaagacagac tgaaactaat agctaatagt 2700

ttttgattac ttggtatatg tccagccctg ttctaactgc tttatagata attatgcatt 2760

gattctgtga aacactctta taaggaaggt actgttatct ccattttata ggcaaggcaa 2820

ctgaggcaca gagaggttac attggtggta gaatacggaa caacattatg aagctagaca 2880

atctgtcatc agatcccagg ttctttttt ttttttctt tttttgagac ggaatctcgc 2940

tctgtcaccc aggctggagt gcagtggcgc gatctcggct cactgtaacc tctgcctctt 3000

gggttcaagc aattctcctg cctcagcctc ccaagtagct gggattacag gcatgtgctg 3060

atcccaggct cctaaccact aggcacacga catgtacatg agtcccctct cctgaatctc 3120

cacttccaga tttagtcctt cactcaggac ctggttcctg atgcctgcag ggtttttagt 3180

gttttgctgc agtcgggctc tgggacagtg tagccttagg gcatagagca aggagagagc 3240

ccctccaatg aatctgcagg agggccagga taggaaagga gggagggcag caaaagcaaa 3300

tacctggaag cagctagtaa gataacagaa cagcgagaat cctctcctcc ccaggcaata 3360

cacctgcatc ccccacctct cccaagtaca cagatatgca tgcactcgca tgcacataca 3420

tctgtgctca attattccag tggaagggca gagcagggct gccagagaag agattgcttg 3480

tcagtgaagg agtttggcag atgtgccctc cagctctgtt tttgaaagaa gccagagctt 3540

tgggggctgg gttacctgca gaactccagg agcttagaga gtgcctccag agagggcaga 3600

tgctgagagg acctcgggct gcgctgcagg ataggctcat tcaaaactct gtgagccaag 3660

ctttgtggac agcctcctgg actcttcttg atcaagttct aagactccat tctggaacct 3720

tcctaaggca ggcttagtgg caatttttct cttggccctg attattctgc agaaacccct 3780

ccatcttctc aatcagtggg ctacctgaaa acagaactca gtcttcatca gttccttagg 3840

ctttgactcc tctgcataaa aatagtcttg gtgacccctc agttttccaa gaataagctg 3900

gctaaacatg aaagcaaggt cttcccatag tgcttctctc ccaaaagctc actgtactgt 3960

gtggttagaa cctgtctacc tcacatcatc agacatctct gcaattggca cctactcttc 4020

tatttttatt cttactttta caggctcgcg tgtacttatg cgactcagtt cactaaaata 4080

aggatggtac tgtcctgtat atctcctaga agttttgaaa atctgatgag ataaagtcct 4140

gtatgccctt gtgaagtaca cgaagtgctc tatgtgctag aaaacacatc catctcaccc 4200

accagcctgt tagcttcttg tctgagtgtt cccttcctag tgcccaacag agtactttgt 4260

acatagtaga ctcccaataa aatggctctt cattaattca agatatgact gggaggaagt 4320

gcagcatttt aggattgcaa atgactatca agcctctgtt actgtgagtt gcttcatggc 4380

ccacgccttg gactaggccc tacccgtagt ccccaagagc atcatcagat gacacaagtg 4440

aagccaactt cacaactttt tctttagaag ggtacactgg gagtagggct atgctgggct 4500

atgaggagat tgcgttaagg ccctgtctct cccctcatta ataacaagga ggtgggaatg 4560

agagcaatac ctgggaccaa cactgtgatg gtagaaggtg agggaagaca gaagaaggca 4620

tggggacttt gggactctag cttcttgttc attccaaact atttcagatc cattcgatgt 4680

aagcagaatg aagggagaag gctgagtggg accatctccc tctagcagca ggcagacagc 4740

ctaagactca gagaggggaa atgatcacag ccccatagac ctgcaggagc cctttgaggc 4800

cacctccttc agtttatagt tgaggcccag aaacctgaag taattttatc caaggtcaag 4860

aaccagggaa ggacacagcc atgactggaa cccaggcagc ccagctcata gatcaggctt 4920

tgattcccta ccctcacaca aagggaaatg gttctttatt attccatatg ccttaaccag 4980

tttgaccttc acagcaagcc cagtagtcag gcaaagggcc atcattccca ttttatacca 5040

gaggaatctg gggcccaaac agatttttc actaatagta catattgggc caggtgtggt 5100

ggctcacgcc tgtaatccca gcactttggg aggctgaggc gggtggatca cctgaggtca 5160

ggaatttgag accagcctga ccaacatgga gaaaccccat ctctactaaa aatacaaaat 5220

tacctgggcg tgatggtgca tgcctgtaat cccagctact cgggaggctg aggcaggaga 5280

atcacttgaa cccaggagat ggagattgca gtgagctgag attgcaccat tgcactccag 5340

cctgggcaac aagagggaaa ctccgtctca aaaaataaat aaataaataa ataaataaaa 5400

tagcatatat tggtgacatg gccaggaaca ttaactagtg ctcttaacag agcagctgtt 5460

aagggcatta cagctagtca gccaaacctg ggtttgaatc ttggctcaac attattactg 5520

tctgtgagac cctgagggaa gttgcttatc ctctttgggg ctcagtattc tcaactataa 5580

actagggaac gataatatta ctaacctcac agcctgttgt gagcacgaag taaagcaata 5640

tatgtagagt tcctgacatg gttggaggca tgtagtgagt tctccataac acagtaattc 5700

ttcctttctc tctgcaccca ccccaacctc agtgcctctc tcttcaggag tccaggagga 5760

tggggagggg cagttggtaa tagagttgaa gaagtatgct ttgggacaca taaccagcaa 5820

tattccctta tcacctggcc ctagctcctc ctgcctccct ccctcatccc accaaggaca 5880

ttggacgaaa gagtatccag agcgcactgg tactccattt gtgttcctga gctggcagca 5940

tgcaaaacct ttaaatacaa attttaaatc agcaaatccc aattgcctga tcttggaggc 6000

aaaaaagaca gatagacaac agatgtagag aaaggagcca ggaggaccaa ataccccaca 6060

gcttggggtt gggggaggag gtagagataa agacgaaatg aaggcgagga gacaagtgca 6120

cacaaatagg gatgcttacc atgtggccga attccttttc ccctaggcag acagtcaagc 6180

agctttgtgg cctggtagga ccaggcctcc ttgtttctcg gtattgctaa aaccaaggtt 6240

aatgtcaggt cctgacactg gcccaatgac aatcctagcc agcaggagct tggctcagaa 6300

gaagtgaaaa aaaaatgata agcaagtgct ctttgcagaa gtctttccct tttatgaccc 6360

catatatatt tgtctgttcc tttgtcatct ataccatttt cttagtttgt agaaaactgg 6420

cctgcttctg tcttgggacc actgcaaggg aatgttattc cccagcaata aagtaggtca 6480

gagccagact ggagacttgc attcctgtgt ggttgtctcc attgtcacca ctcaagtctc 6540

agggattgct cttcctaaaa ccatgtcctt gggagcaagg ctttgcccat tgttttctct 6600

tctcttctct tctctagcag aaggtctcta aaggctacta tccctagaac tggccatctt 6660

caggctgaag ctaggttggg ctaaaatgat cagccttcag tagaaatcaa ggctgaaaca 6720

ggtaaaaagc aggctgtctg agcatttccc taccagagac ccttcagcta agtgtctggg 6780

tatactcagt caaacaccat ccctgccttt gtgttcatat tctgtgtccc tattcaagaa 6840

aggtagtttt ctttctttct ctctttcttt cttttctttc tttccttttc tttctttctt 6900

ttctctctct ctctttcttc ccgtctttct ttctctctct ttctttcttt cctttctttc 6960

ttttttttt ttttgagac ggagttttgc tcttgttgcc caggctggag tgcaactgcg 7020

caatctcggc taaccacaac ctctgcctcc cgggttcaag ccattctccc gcctcagcct 7080

ccggggtagc tgggattaca ggcatgcacc atcatgcccg gataattctg tatttttag 7140

tagagacggg atttctccat attggtctgg ctggtctcga attccaaacc tcaggtatcc 7200

acccgcctca gcctcccaaa gtagtgggat tacaggcgtg agccaccatg cccggcctgg 7260

cagtttcctt aaaccagctg tcaggtgtct cggccctcag cttcctgagc tatgctgccc 7320

atttctcaga gcagccaggg gaatctctgg gcaagtgtag tgttgcatct gtcagaagaa 7380

gaaataaagc tttttaatga tccttgagtg aatagagagg ttagtttcat taaatacagg 7440

actgccaagc tggcagcctt actgattagg ctccttggat caagttctca gcacaaagat 7500

ggtggaattg ggattaagag gaacaaaggg agcccacatc ccaatgagaa acattcagga 7560

tgtggaggcc ccttgctggt ctccaatgag gtgtttttgt ttcacttttg gtgcccttca 7620

gcaggggcac ctcctccagg tggaccaatt gacatccaag gacacgttta aatgttcagg 7680

tgccatggac tcgatgtgtt ccttaaaatt catatgctgg aaattaatct ccaaagtgac 7740

agtgttggaa tatgggcctt ttgggaggct ccaaccacat gaatggatta atgtcaacta 7800

taaaaagggc ttgtgggagt cggttcactc tcaagtcctt ctgccttatc cgtttcaatg 7860

ctggggacac agcaaggccc catcttggaa ccagagacca aatctattgg cgccttgatc 7920

ctgaactctc cagcttctag aactatggaa ataaatttat gttctttata aattacctag 7980

tctttggttt tctgttatag cagcacaaaa cagactaaga caacagatgt ttagtttctg 8040

ggctcttagt attccctgct cactaccact tccctctcaa ggaagaagaa gggtgaagca 8100

ggaaaatgga aacttgcatt tgctgagcaa gtattggttg ccagacactg ctttgggcac 8160

tttcctatac cttatctcat ttctttcttt ttttttttct ttcagggtaa gggatgtgct 8220

tcaggtgcaa agaatttcat tgactggaat actgttagag tcatattatc atcatcagtt 8280

tatgcatggc ccactttgta tcattttatt tctattttta tttatttta ctttatttat 8340

ttccctcttt atcacccttc ccaatcccat ctcctttcat gatggatatg cactcttttt 8400
tttgagacag aatttcacgc ttgttgccca ggctgcagtg caatgatgca atcttggctc 8460
actgcaacct ccgcctccca ggttcaagcg attctcctgt cttagcctcc agagtagctg 8520
ggattacagg catccaccac catgcatggc taatttttg tatttttaat agagacaggg 8580
tttcaccatg ttggccaggc tggtctcaaa ctcctgacct caggtgatcc acccgccttg 8640
gcctcccaaa gtgctgggat tacaggcgtg agccactgca cccggccaga tatgcactct 8700
tgtgtgtttg atttacataa atggtattgt tctgtacatc ttctctgctt ctttgtttca 8760
tgtaatatta taattattga gagctaattc tgttgataac tgtgaaagta atatgttcat 8820
tttaaccact atataatatt ccatgatttg catattttgc attttattta tccattcctc 8880
tgttgatgga catttaggtt gttttcaatt tctctgtatt acaatattgc aacaaacatt 8940
tcatgcaagc tactttgtgt acctgtgaga gtagttctag ggggatatag aaatgttggc 9000
caggcatggt ggctcacgcc tgtaatccca gcactttggg agccgagcgg cggatcacga 9060
ggtcaggagt tcgaaaccag cctgaccaac atggagaaac cctgtctcta ccaaaaattt 9120
agtagagacg gggtttctcc acgttggcca ggctcgtgta ggactcttga cgtactgatc 9180
cgcgtgcctc ggcgtcccaa agtgctggga ttacaggcgt gagccaccgc gcccggcctc 9240
tccaatattt ctatagttgt atagttttgc ctttcatgtt taagtgtgca gtacctctaa 9300
tgtttattac gtacgtggca tgacataagc atttaatctt gtttttctat gttataagcc 9360
aatttctcct attttatgcc tgttattcat agcatggacc tttcctgtgg ctcttgcttc 9420
atgagcagta aatcaagata cattcaccaa gttcccatcg tattcagagc cctgaagaag 9480
cgggattcta gtggagttta tgacacagaa cctgccttca ggaaatatgc catcttgctg 9540
ggaaaataaa cccatggaga gaagaaaagg ttttgagagt atacataaag caatacctaa 9600
gtgaggataa atagaaacct tggaaactga gaatacatgt ttggagggtc aaaaagggag 9660
cagggtaccc attcctaggc tgggatggat aaaaggattg aactagaagt acagggacac 9720
cattggatat tttcactgtg agatgatcct cgaatcaagg tagctttatg cagttgccac 9780
agaatgaaag aggagatagg agacactgag acttcaccat tggccattat ctggccttgc 9840
tggtccgctc agagtttcct tcccatgccc ttctcctttc ttattgctcc tcagccccag 9900
attactaatt aataaggaag caaggtcatt ttttagtgca gttcaggtac ccagtccaaa 9960
aaatctccct gaatcagctg tctccattct cagagccatg gattaaattg aagggaacta 10020
atttaaaata ataataataa cacaaaaaca ctaataacaa caggaataat tgtaacataa 10080
taacctcaca ccctttcagg agctcatatc ttctctatga atatgatcta ccttgctcac 10140

cagtgggaag agaggaagct gattgaagga gattttgtgg agcgggtgcc tggactctgg 10200
tagaaaggga tcttggggtg tggcagacac aagacagaag gtcaggactt agctgccttt 10260
ctcagcctaa actggctttt agtttgcctg tgactttcct atttggccag aaaggaaaaa 10320
gtgtggaggc ttcagagagc ttcaacgccc cctcctcccc tagcaggccc cctctgcctc 10380
aggatgctgt gccatgccac aaaggtggta catcttgcct tccaaacttg gctctcccag 10440
agtctacatt acaaagtgta caccaattac cctcagcagg gcagcctgga tttagtgtag 10500
cccacgtcag gagctgattt tgctcctagt aatctccttc tgctttagtg gtttgagact 10560
ctggaggggt aggtgggagg ggaacagggg ggcaggtggg tggaggaaga aaagcaaaca 10620
aaactcatca aaactaacaa acaaaaaaac tccagctctc tgctcattaa tcctctcagg 10680
gcctacttct taccatggca accagtaagc ctccaggcct cctcagtcct cactccagtt 10740
gttagtaaag agaggaggct ggctgctttg cctgggatga tttcaatact ggacaaatta 10800
aactggtcct ggggagctgt ggaagttgaa ggaggaaagc ggtagagaga ggaaagggaa 10860
tgctgaagtg aggaatggca agtttccatt aggcagcatg tctcaggatc tctcttggga 10920
aggtggtaag ggcctccaat gccaacttgt gttaaccctc aacggtcttg atgaaggcca 10980
accctgccca ctctcttccc cattccccaa aggcaagatg atgcagtgtg tgggggtgtc 11040
agtcagctct ttgcagcaca aagtgatttg ctccgtcttg aaccaaaggc aagaagagga 11100
gacagaaagt caaagtcagc agggtactta ctttcacttc acccagaagg cagctggtca 11160
ttggtgagag tgcaactctc taaagaaaca gaagacacat tcctcaggaa aggttcaggg 11220
caacttagct tctctgacct ccaactctgc ctcaccaagc ctcccaaagg caccattgcc 11280
acagccctca tttggttcaa tctgatcaca ggtcactgag ccccatggat ggataaccat 11340
ctgccaaggt gggcagtttt tcaagttata tcagaggctg tgtaactccc actggcttct 11400
ttcctaagaa actgaggctg gagccagatt ccttgaccca caatgaactg ggacttgact 11460
atgatatcac tcctgaggct agtttccagg caattcacat tggtcctaca acacggaggt 11520
atgaatgaat tcttaggtcc tttcttaaat tcctgctttg acaaggaaaa ggcagaatgt 11580
tggaaaggaa ggactcaggt gaggcccctt ggtgtttttg tcctgtccac agctagtgag 11640
tggttatttg ttcattttgt ttccttattt ggtttggttt ttcctattac aaagggaagt 11700
agtactactt tctccctaaa aacactatct ttcttcctct gtatgtcctc ctgtgcctga 11760
tactatctgc acactttccc attcctttag tgagtacctg agtcccagag tctaaaaaaa 11820
aatcatctcc gcttcagcct cagcctccaa gaccttaata gcttaggcat attaatttct 11880
gagccagaat agagtcttgc tttcacttac tcccaaatgc ttgttgaaag tcaggtaatt 11940
tggctaccac tgtatctcta gtattccttg gcacatggtt ggcattcagt aaagatcctt 12000

tttttttttt tttttttga gacagggtct gactgtgtcg ctcaggctgg ggagtgcagt 12060
ggtgcaatca tggcttactg cagccttgac cttctctgct caagtgatcc tcccacctca 12120
gcctccgaag tagttaggtg ggactgcaga tgctcacccc cacacctggc taatttttt 12180
aatttttttg tacagatggg gtctcattat attgcccaga ctatagtgaa gaaattttaa 12240
acgaatctta gtatatgtat aatctcagag tttttcgctc acttaaaata acagcagcat 12300
ctaaagcttc catttattga aaagctactg tgtcagacag agcctgtgct atggactaca 12360
tatggacatt ctaatttaaa acttaggaca tttttgaggg acagaaactg ttattaattc 12420
tgttccacag atgacaatca tgaggctcaa gaatgttgat taattttccc acatagttaa 12480
taaatggcag aacctagtac caaactcaga taccttggac tccaaagatt gcctctctat 12540
gcctctggtg ataggaaagt gccattgaga ctaagatcac tatttggctt tctgacatgc 12600
cagttaggtt tgcacagaaa cagttccatg gctacatact ttgttctcaa ccctagcaaa 12660
tcagtttatc caaaggggct caggtgagag tgtgctaagt agagagaagt tgccaagtac 12720
ccttttctac ttgtaaatat cccttacata tgtacaaagt ttccgtttaa aaaaatactt 12780
ttttatttcc tacagtcttc acagcaagca tgagagagtt ttgttacagc tactcagttt 12840
ttaaataagg acactgaaga tcagggaagt gttataaatt gtctaaggtc acacagccaa 12900
gaagtagtag aacaggggct aaaatccact cttcttattc tagagcacca gagtataagc 12960
cactggatcc ttttggcaca taaagcatta gaaaccttgc tttgtgtcag atgccaactc 13020
ccacaaaaag attactgctt gctggtggcc tatgcctgta ttcctggcta cctaggaggc 13080
tgaggggagg atatcttgag cccaggaatt tgagatcagc cagagcaaca cagaaagatt 13140
ccatctaaaa agactgctgc attttattca gacaaggggg aactagagcc ccagttctat 13200
gtcccaaatg gctaactgac aattgagaag aaggcttgca tctcatatgt tctgtcagaa 13260
ttcgaaagac atcatcagga aaacctttt atatcttctg aaagttgctg tctaagtatt 13320
agcaaccccc ctgccacccc acaacaggca ttgaatgagg actgggaatg cacatcactc 13380
ccctccatct gtcctctttc ccctttgagc agtcttaatc attttgaaat ttcaatgaat 13440
taagaaaaca ggccatcacc cctagatata acaacacctc tctttggaag gaggcattgg 13500
tcataaagag atcactgctt tatggggcaa aagccatcag attactgggg attttgcacc 13560
agcctaagta ggcttaactc ctaaatggca aagaggcttt gacaaggatc cctcagaatt 13620
ctgctagtga ttcatgtggg gctttcctcc tcctcacctc tctctacttc tgtcacccca 13680
gcccccactc ctcccataat aagcagtcct tccctgcctc tcccactgtc aagggactct 13740
tctgtgaaag aataacattc ccagccgggc aagatggctc aagcctataa tcccagcact 13800

ttcggaggcc gagacgagtg gatcacctga ggtcaggagt tcgagaccag cctggcccac 13860
atggtgaaac cccgtctcta ctaaaaatac aaaaattagc agggcgtggt ggtgggtgcc 13920
tgtagtccca gctactcggg aggctgaggc aggagaatca cttgaatccg tgaggtggag 13980
gttgcagtga gctgagatcg tgccactgca ctccagcctg ggtgacagag cgagactcca 14040
tctcaaaaaa aaaaaaaaaa aaaaaaggca gggcgtggtg gctcatgcct gtaatcccac 14100
cagtttggga ggccgaggca ggcggaccac ctgaggtcag gagttcgaga acagcctgac 14160
caacttggag aaacccagcc tctactaaag atacaaaatt agccgggcat ggaggctcat 14220
gcctgtaatc ccagctactc gggaggctga ggcaagagaa tcacttgaac ctgggagcag 14280
agactgcggc gagccaagat tgcgccatta cactccagcc tgggcagcaa gagtaaaact 14340
ctgtctcaaa aaaaaaaaaa aaaaaaaaag aataacattc cttcttcact gaattattca 14400
tctttgctgc cttggtatgg caaaacttct cattaaaagc cctgggtgag gccgggcgcg 14460
gtggctcatg cctgtaatcc cagcactttg ggagcccaag gcaggcggat cacaaggtca 14520
ggagtttgag accagcctgg ccaatatggt gaaaccccgt ctctactaaa aatacaaaaa 14580
ttagcaggag gtgatggcag gcgcctgtag tcccagctac tcgaaaggct gaggcaggat 14640
aatcacttga ctccgggagg cggaggttgc agtgagttca ggccactgca ctccagcctg 14700
agcaacagag caagactctg tctcaaaaaa aaaaaaaaaa aaagccctca gtgggagtga 14760
cctacttagc gcagcactca caattagaca gagagaagac aaggggagga agagatgcag 14820
tggcaggggt gggaggctac cccaaagggg ttgtgacagg gagtccccaa aacccaggga 14880
cttgagagac tgaatggcac ctctggaggc tcctggctgc ggggatgacc acggtgcgct 14940
gcttaccttt cattcttcat tcacttgtac acatgagagg cggcacacat gctgctgcca 15000
tgaagaccat gggctctggg gccaggctgc tgggtttaca tcccacatat cttccgagat 15060
atgtgaccct gaaaaaccat ttaaactagc catggtggtt catgcctgta atcccaacac 15120
tttggaaggc agaggcagga ggatcacttg agcccaggag ttctagagca gcctgggcaa 15180
catagtgaca ccccagccct aaaagaaaac aaaacaaaaa atttaattag ccgggtgtgg 15240
tggcaagtgc ctgtagtctc agctactcgg gaggctgagg tgggaggatc gcttgagcca 15300
ggaaggttga gactgcagtg agctgtgatc acaccacggc actccagcct gggcgacaga 15360
atgagacact atctcaaaaa aataataaaa taaaagtaag aggaagagac tagggtgcac 15420
gcgctctctc tctcttacag aaagaatggg ggcctggagt ggcaaaacag gttaatgaga 15480
gaaagaagag gaggcctggc cagcaaatgc agtcttgtta tgtaaatgaa acctcacaag 15540
tagcagctct cacagagaat aggtggtaaa tgttcctttc agacctttaa agttgtcaga 15600
atctcagtta atcgttccta gatcccaaaa aggaaaggcc tttggaaagt ctggctgcag 15660

cgatgcaggt tttctctaca aatgcaaatc tcccctacag aagacagctt tgcaaggcta 15720
cttgagtttg caggccctct gagcagtcat ttcaaaatat gtcaaaaaat tatattttgg 15780
ggtaaaatat tttgatttcc ttcactagaa agagggccct cagcagatcc tgaccatctt 15840
ggcaccctgg tcttggattt ccaacctcca gaactgtgag aactaaattt cgcttatttt 15900
ttagtcaccc agtcagtggt attttttttat ggcctcccta gctgactaag acaatgatgc 15960
tgcaagaaac cttgccacca gagcagctaa gatgagcccc catggagcaa ctgacactca 16020
gcactgttgt ccaagctagg gcaggcacca gtgcctcctc atccattctt gtatttattc 16080
attcatgtat tcatttacaa atgtttagct ggcacctacc ctgtgcctgg ccctaagcta 16140
ggtgctgtag ggtgcacagc aatgagaaaa tcaggcccag gcctcaccct cgtggagctt 16200
agaagctact aaggaggagc atctctgaag aatcacacac atgaatgtga gacacaatag 16260
agttttttaa aaacccaaga tgttgtgagc accaataaga ttgggtaagg ggccaggaaa 16320
aaactccaag gcagctgggt gaggtcgctc atgcctgtaa tcccagcatt ttgggagcct 16380
gaggtgggtg gatcacatga ggtcaggagt tggagaccag cctggccaac atggcaaaac 16440
cccgtctcta caaaaataca aaaatttgct aggcctggtg gcaggcgcct ataaacccag 16500
ctactcagga gcctgaggca gaattgcttg agcccgggag gcagaggttg cagtgagcca 16560
aggtgcacca ctgcactcca gcctgggtga cagaaggaga ctccatgtca aaaaacacat 16620
aaataaataa gaacgccaaa gagatgatgt ttaaatctgg acactgagaa agaaagagaa 16680
ggcatcaaac aggcaaagag tgaaggagaa gttggtgcag gtggaagggc aaggtgagtt 16740
cctttccctg gcccaggtca tttccaaacg gttttgtctt agaggagtga ggctctctga 16800
ggtctggatg cttattggct tccacttgtc aaactctgac tgccaccaca tgggctttcc 16860
caaagcctag cccacatacg acgtgtatca gagtcaccag gaaaacttga taagatacag 16920
atacctaggc tgcactcgcc tgagagcatt ttagtgagca gtcatagtgg ggcccaagaa 16980
tggtggggcc caagcagtaa tgatggggcc caagaatgca ttttaacaag gaccccaaga 17040
ggttgctcaa gcttggagaa tttctttttg ttttttgttt ttttttgag gcagagtctc 17100
gtctgtcacc caggctggag tgcagtggtg caattttggc tcactgcaac ctctgcccca 17160
ctgggttcaa gtgattctcc tgcctcagcc tcccgagtag ctgggactac aggcatgcac 17220
caccacgccc agctaatttt tgtattttta gtagggatgg ggtttcacca tgttggccag 17280
gctggtctca aactcctgac cttaggtgat ccacctgcct tggcctcccg aagtgctggg 17340
actgcaggca tgacccaccg tgcctggccg ggaatctctt gacatactac atagtatgag 17400
gaaaatgggt tggtggcaaa aactgagcta gtaaaccacc gcccatgttt acaagcctgg 17460

gtttaatagt gactttctct caaaagctat ctctggtctc tgcatgtcgt tttttctaat 17520
atcaggctcg gaagagtgta ggctctcttt agcttggagt ttctttttt ttttttttt 17580
ttaagacaga gtttcactcc tgttgcccag gctggagtgc aatggtgcaa tctcggctca 17640
ccacaacctc cgcctcccag gttcaagaga ttcttctgcc tcagcctccc aagtagctgg 17700
gattactgcc atgcgccacc acgtccggct aattgtgttt gttttgtttg tttgtatttt 17760
tagtagagaa ggggtttctc catgttggtc aggctggtct caaactcctg acctcaggtg 17820
atccacccgc ctcggcctcc caaagtgctg ggattatagg cttaagccac cgcgcccagc 17880
cagcttggaa tttctaagtg gatttttttct taaggagacc aaccctcctg gtttgcctgg 17940
tactctccca gttttagcat taaaaattct gtgtcccagg aaatcctcag tcttgatatt 18000
aattggatcc ctaatatctt tggcaaatga ggatgactgg ttaacctagt gtcctagaga 18060
agtgcctagc tggattcacc tagaaaaggg aggagtcact aaaagtctcc ccagtctccc 18120
tactgccaag gtgctcccac ctgccaaact tgccttctca tccatgtgac aagaggtaca 18180
cagaagagtc ttctccctcc aactccccag catgcatgct ttctcctgtt tgtttctgga 18240
cagtgtcata tcaaaaagga gacatgacaa taagccagta ggtcccacag cataaccaat 18300
gtgtttgaga tgggaatgat caaagcttta ccaagggcta agccctggtc ttagaaactg 18360
gagcagaaag aaactatcag ctaacatatt cacagggtaa tagaaaagat gacatatagc 18420
atcgccatgg caaccagcct ccaaaggccc ccagcacctg ctccattcta cacaatattt 18480
ccccaaacag ggatatttct tacttccttg tttttctttt caataaagat gttaatttta 18540
ttccccctca tttgtgctta atgttccagt agcaatttag ctccttctgg gctttcctag 18600
cccctccctg tctctcccaa acctttttctt ctcctgtctc ccatccttcc tttcccttcg 18660
atctcttttt tttctctgtc cccacaattc ctgtccctgc ttccctagac aaatatttcc 18720
aaccacctct cttctcccag agctaagccc tcttccagga tattttatca gctgcggttc 18780
tgatggccta gtaagtgttg gtctttgcat tgggtgagtt ctctatggtg ctctcttgaa 18840
atcccacgcc ctctcttctg gtgtctcacc acaccaccac catcccatct ccctgagcta 18900
ggggacacag tttagtaaaa ccattttagc ccacagagca gatgaaggga aacaaaggct 18960
gcaaggtgga gaatagtaac agttggggtt ttttgagtgt tgatcatatg cctgacccta 19020
ctttagagct ttatatctat tgtcacattt aatcttcatg acaattctga tggcataggt 19080
tttattgact gattgattga ttgagacaga ctctcgatcc atcacccagg ctggagtaca 19140
gtggcccgat cttagctcac tacaacctct gcctcccggg ttcaagcgat tctcctgctt 19200
cagcctccca agtagctgag actacaggca ccagctgcca catcctgcta atttttgtat 19260
ttttagtaga gatggggttt caccatgttg gccaggcttg tctcgatctc ccgatctcaa 19320

gtgatccccc tgcctcggcc ttgtaggtat tatttatgac ttcacttcac acatagggaa 19380
acggagactt agactggtta tgacacttgg aaaaggtcat ggagctagtt agtagtagtg 19440
tcaggattta aacctaggca gtctgacagc caagtctgcc ccctttcatg aagaggtttc 19500
aaaagaagat cctgtctgca ttcattgaat gggaatgggg tcctggcaaa ccctttcatc 19560
agcattattt gaagggtgcc cctggtcaat aaaactgtcc tgtaaaaaaa acacatgtcc 19620
ttctgcttag gtccatggga gcagtgaagc agaaagggag tggaggagct ctcaccctga 19680
gcctcttctc tcttcagttc gttcacacag tctctcattt gttcattcat gcaacacacg 19740
tttagcaagt atttcccatg agttaggctc tgcccagcct atccatagga aactgtgttg 19800
tggagagacc gctgagtgtc aggaatcctg tctatgaagt ctgtctctcc acttgccatt 19860
tctgagcttg tttctgggga caatgatgct tacgagttga cagaaatcac ggctatggaa 19920
acacctaaca tgatgcctga catagagggg ttaatttttt agaggcagaa tcacttctac 19980
atgggaaagt tgtaccagaa acaaagcagg gtggtgctga attttcccag aggtgagatg 20040
ggcctgaagg tgcttggctc ctaccaactc tcagggttgg ctctctcaaa gcagggaaac 20100
ccaaggctgc ccatggcatg cccctactta tctcttagag taagcagctt cctctttcca 20160
gatggtttac ttcttgtggt agaaaaagag acccagccct ctgcaaagca agaatttggt 20220
tccaaaacca atcttctggg ctgacaaagt ctctagatac agacagacac tgctgaggca 20280
acattattat tattatattg agacaatgta tgcccagggt tgcaagggaa ggtggttcca 20340
cagaacagac attaaataaa ttatttgggt aagggaggag agtacaaaca attttaagga 20400
aatctggcca tttccatttc tctctcttgc tttctcctct cttatcctct ttattctctc 20460
agctctgatt taagaaggag aatgagttac tgagtaaata gtggaggaag tagtttgtgt 20520
gtggatactt tagctgactt ttggcttaat attaaaggac ccccttggcc ccactgctga 20580
acattttctg caacttggga aggccagatt gcctgggcct gtgcctgtca gggagagttc 20640
atggctacac ggtgcagtgc actgtgtata ggagtgcaca tgggtgcctt gtggagagag 20700
gccaagggga aatgtcacaa agcacaaagc agaggacagg tttgagcttt tgggggaaag 20760
cagtaatgtg acgggggcag ctggcctgga agccaaagtt tctgagatct aatctaaacc 20820
ctgcaccaaa tcacggatca ccttgagcct gtcatttcac ccttctgtgg ccttcaaaaa 20880
gaggagactg aattagatct ttaatatctg gtcatttcta acgttatata acttggtggt 20940
tttgcatgag atgttgggga gtatggggga ttctagccct tcttgttcct ttgccccatg 21000
caagaatctc cacttgaatt tcagtctttt tccctttcct caccaaatca atatatttca 21060
ggcaccaaat agaagccaca aggcacttct aaattgtcga aaattacagt caaggctact 21120

tccatctctg ctcatggtag tgctaggcca agctataggg cagggttgaa gagggttggt 21180
ggcaggagtc aggaagtggt ggagcaaaga agatttaacc ccttaaaaaa aaaaaaaacc 21240
accatattca tttcatccat gatttgtcta gtggctgaaa tgatcccctt actaggatgg 21300
agacaacttc accgtaagtc acagcaaagt ttcatttggg aggaagtatt gatcacaact 21360
actcactgtt acttaatatt taaggatcca gtttcttctt acttgactgg agcagccagt 21420
tgggggatag gcctacgtgg ctttgtctct tgacttgatg atgaacttgt ttgactttgg 21480
gaaaaaggcc tatgccatcg ccaagctctg tggaatgcag gctggagaaa gtttagtttc 21540
ttcctacaaa agtctccagg aggtcaggaa cagtggctca cacctgtaat cccagcattt 21600
tgggaggctg aggcaggagg atcacttgaa cccaagtttc aagaccagtc taggcaataa 21660
agtgagaccc ccatctctac acaataaaaa cttaggtatg gtggaatgca cctgtagtcc 21720
cagctactct tgaagccgaa gtgggaggat tgcttgagtc cggaaggtca aggctacagt 21780
gagcctgtgg ctgtaccact gtgctccagc ctgggtgaca gaatgagacc ctgtctcaaa 21840
aaaaaaaaga gtctccaggg ggcctttgat tcctaggaaa gcaggatgct tatcctttct 21900
ctactctcta ctcatctgct gagcccaaag ggagctgcat ggcagaatca aggctcatga 21960
acagatcatt ctatacttcc tattgcccca ttcttcagat gaagagaatg accttgattc 22020
ctattgggta tcaggagact tcagttgatc tctggaagat gatactctga acctgcaggt 22080
gagtagaggt ctgagtagat gcagagcagt ggagaggcca aagactgaat atgaaactgg 22140
acatgagaca cacatctaaa tattttcaga acactaacta caaaaagaag ccttgcatta 22200
tttctgctcc agtctatctt tttggttggc ttgattttca ccaacttgca aacatggtat 22260
tagttatgtt tgggaaagtg gagttggaat gagagtgaaa gtataaaatg tgaggcagag 22320
aagacagaga tgtatagttt ttagtcttct ttttcccatc tgtaaaatga atctaagaaa 22380
gaccgataat agtagatcac attcaaacta ggagttaaag acactgaaat tgattgaata 22440
attgatatgt attccatcaa cttgaaacaa aggattttcc tgatctctga taatttggac 22500
acacatgagg gccagaaact tctctttcat agtaactaga atgactttcc agtcttcata 22560
gggtcaaaag aatctttttg gggtaaagaa atggggaggg acctatggaa ccattagaat 22620
ctatggtgaa ctgagctcta cactaggatt attattatta ttattttgag acagagtctt 22680
gctcttatca accaggctgc agtgcaatgg tgcgatctca gctcactgca acctccacct 22740
cctgggttca agcgattctc ctccctcagt ctcccgagta gctgggatta caggcgcccg 22800
ccaccacgcc cagctaattt ttgtattttt agtagagaca gtggtttcac catgctgacc 22860
agctggtctc gaactcctga cctcagataa tccgcctgtc tcgacttccc aaagtgctgg 22920
gattacaggc gtgagccacc gtgcccggcc tacaccagga tttaattaat gcaattaaca 22980

tgagctaaga ctccccaggc cctggtcttc ttagcactga ttaacagctg ctccaggatt 23040
tggtctctgg tctgaggctt atctctttag tttagccttc tctgggaatc ttctgactgc 23100
tgctggttgc tactttatct tgaacagaat ctcagcaaac atttagccag aatcagagcc 23160
atcagaaaag agccacttca aattctctac ctccagagct caggatgagc tgagctccct 23220
ggccctcact tccagtttcc tcctgggtgt ttctctgcag tgctaaaatg ttcatcccta 23280
ctcagtgaac tttctaaata gtcctcttct actttatcct accatttcat atacctggtt 23340
ctgtacacgg aaaataattt ggtctcccaa gtgcttttgc ttaaggtgat gcaactgagc 23400
agacttcctg gatctatgcc atggcttgtc tgaaaatgtc aagtcctcag catcctattt 23460
gtctttcctg aaactggccc ataattgagg ctgattgtac tggatttgca gttgtgaaac 23520
agcacctgcc tttgactcag cagtgggcac actgctgtgg aaagtgaggg gattatagaa 23580
cacgtgggat ttaatgtgtg gttagaaagg caatgtaaag gttacagaaa gaaactgaca 23640
atgggcagga aagaaaagtc agagttctca aagcaacact ctcttagtgt ctgacgatga 23700
tgttgctgac agtgcaggag tgggtggcag caaaagctca tgtgtgtgac ctgctatgtg 23760
gtttggagtt gcctattgtc tgattagtat ctgagatgac tctttaggag tgcctgtcat 23820
tgattagaat ctgtgatttg gggccatatg cgacataact ctgcaaaatg aatacctccc 23880
ctaccttgta caaagagaag tggctcaatt tggcatctta cctagggatt ccaattattc 23940
tgggcaagtg actggctaga agcacttcat tagcacagta aatgtagaga cctcagccac 24000
aaaacaaaga caagatctac ctgtactcag gataaaatgg actagtcatc acactccaat 24060
ttctttggcc acacactctt ggataataat tgttagcaac agggatatca tcttcagatt 24120
atcaaggaga gttctgtgtg tgtgtgtgtg tctgcatgca cacgtgtata ctcagataat 24180
gaacactaca gaagtgcaaa tttaggctag tacatcattt gtaatgttaa atttataatt 24240
ttggcatata acatttgcca actataacta tgtgattggt gttagtaatg caaattcaga 24300
tttgtgcaca attctatgta gtaatcaacc aaactgttct ttactttagt gtccagtgtc 24360
ctattacaac aaacattaaa tttattcata gtaaaaccaa aaaagaaaaa cagaagaaaa 24420
gacaggagaa cagacagaaa gatagactga aggaaagata aagtgaaaag attgaagggg 24480
gaagaaggaa ggaaggaagg atcctaaata attggaaacc actcagagaa gcaccacagc 24540
taaggatgct gttaatctcg cctaactctt ctagtgagga aggtgcttat taacagttgc 24600
agacaaatat gtatgcatca tgagataatc atgagattat cttttcaggg ttcttgggca 24660
tgcgtgttct gtagggttgg attattgaag gataagcttc ttgaaagggt gtctagcagg 24720
tgatgacctg tgtgatgtga ctagcagaaa gattctggga tagataaaaa tgaagctcat 24780

tactgcctgc tgaaaaaaat agtctatttt ctaggccata taataaatgc cctactcata 24840

tgctgtgaga ggcactcagg ccctggggca ttgctgatga ctcaggcaga agccttcttt 24900

gggaaggaat tacattagtg acccactaga atctaagcta catgagggca tggactccat 24960

ctgctttttt cactgttgta tctccagagc ctagaatatc tggcatctgg catatgattt 25020

gcactcaata aatatttgtt gaatgcatgc atgaattagc gggagtgtga tccattccac 25080

acagtaaata ttcatttttc caagttctgt tgcttttatg ttatgtaagg gcttcatgcc 25140

ttacatagca tgcattatca caatccccaa atgagtaaaa taagatattc atgaacatga 25200

tcacaagcaa aaaccaattt acctttgagg agaggcatga catggtccta tcatccccta 25260

acaagaatgt tgcagagtta caagagccaa gagacaagga aagtcaatgc aaactctgtg 25320

gaaaccttaa tgaaactctt ctgcacagct agtgaattgc tctcaagtgg ctgatgtgat 25380

gagtaacgat tctgaaggtc cattcagaag gagagaaatg atgcatatct ggatgctagt 25440

cactacacct tccctagctg gagcttcacc actctgttgt tccctggatt cttaacagag 25500

agcccagcaa cagaggcagg gccattgttc actggtgcca tcctggtgct cacatctggt 25560

ttgtctttgg agactgagag gaattcggtt aacctctcca ctggcctccc caagttcaca 25620

tacttcacgt ttttgtctgc aaaactcata cctgaaatat tctattaaac aaaagaatga 25680

agccaaatct ttgaagcatc agacttgaac cagtcactgt ccaaaagtca agagagcact 25740

gaatgtgtct atcatatctg caattttgtt cttccttact ccacctgagt ctaggactgt 25800

gaaacctcca ccaccactac tgcccccaca cactgggcac catcccctta attcccaagg 25860

gtggaagagc ttcttgcctc tgatgccagt gctcctcagc accagaagtt gagcctccgc 25920

tttcaagtaa attgtcatat ctttggtagt tgaaaataac cctgaaaaga aaaattcaca 25980

cgcagcaaaa agatgcaggc ttgctcctta gagatccaac tgaaacgcca cggtcattgt 26040

cagcagattt ttaaaaaggt gtctgtcttt ttttcaaccc ccctccaact acttagccta 26100

gtgcttctga agagtagcta aagctctgag agtttcctca gattggggaa cacagcctgc 26160

ctttcaaggt acctctcagc tctctctcta cttcaataag tgagataggc aatccggggg 26220

ccataccatc tgggtaattg aagttgagtc cagaaaggta tgccaaatgg ttctccaact 26280

ttgcccccca aaatgtagcc ctagaagtca ctgaatttac caaacaatac ccatcccaat 26340

gcagacctgc tgctcctcag actgaggttt caggaagcct gcttatcagc agggcacagt 26400

tgcaattttc cctgcatagg gttgaaacgt ggcagactct tttctggaat gaaacactta 26460

cgaaaatgtg cttaatccat ttttagcagt ttggtagttt tttcatgaca aaataatatc 26520

tttcattctg aaaaacattt tgcctattgt agcttctata caacttcagc actgcaataa 26580

agccagttct cgggtgagta tggcagccag ttgcaaaaca ggaaaagggt gcgatacatt 26640

ttggaacatg atgatgcata tactccatag cttagtgatg agaagagaaa aaggttatgg 26700
ggtgaggctc aacacaacgt tgtactcttc ccacagcaga aaatgctgag gtcatctttg 26760
gaaagtgttg cttcctgttg ttacagggaa ttcttccaat tcccacctct cctgctgttt 26820
gtgaacgcct agtgcagctt acccagttct ttacctccct ttctcccttc tataccccaa 26880
gagcctgctt ccagcaccct aggaaagctg tatctcttca ccatgtcagc caaacatttg 26940
gtcatagcta cagcctctca caagaaggga tctttcctct catgaagctg atcctgctat 27000
gatctcagca ctcagagatc caatctggca cacaggaggt gggaccgaga caatagacag 27060
gtccccaggg gcctatgcct tccattagct aagaattcca gatgaccgag tgtgaggccg 27120
cagacctaag ggacaaggga ataaaaactg ctttctcatt tttccttatt atcattacgt 27180
catgctggcc tctctctgtg tgtcaaactt cttgtttgac agagagagat gcagaggcac 27240
tgagcttgag gaaatattcc tgcccaggat gcatctttta acaaatgttc actaattgcc 27300
tatagtgtgc tggggatttg gtggatgatg tgtctctgct catcagggag cttgcgatag 27360
taggggagac aggcaagtaa tctggtaatg tgacagtgtg acagagtgct gcaaaggtga 27420
gggtcctgtg ggaacacatt ggacagtctc ctatcctata cccaagggac atgggggcag 27480
ctccagaagt gttccttaca aaatgacaac gagactcagt ccagaagggt aagagtggag 27540
ggaagagtac tttggataaa gggaacaaca tacaaaaaca ccaggaagct ggagagacag 27600
aacaggaaac tctagatgca tcagtgggac aagcagagtg accaaccttg ggctgcagcc 27660
tgagagcaat ggggaactgt tgaagcagga gaagtgacat gttgagtttt gccacaggaa 27720
gtgagttaca cagaggagac aagagccaca cccctccccg ccagaccccc agcaaggagc 27780
tctgtcctgg gaggttctgt aaggctcagt ttctgcccat ggagtgggga tcatgatgct 27840
taaaggaaac aaggagggag ggtgagcaga tgaaagggca gctaaagatg ttttaggtta 27900
ttttacattt actgaccata taggcattat ttgagacttg taatgttctc agtcccaact 27960
ccaaaagccc agcacccctg tctatatttt agtgtgcatt taaaattctc ttcttaccct 28020
ttcatttatg gtttgaaaat ttctcttcct taataataac acagccaaaa gggagctgca 28080
gcacaggttc cactcacggc tacctcatta ttaatccttc catatagcca ataggccatg 28140
tggaggagct tccaatcccc tcaaaccctc tctggcaagg ttgaggaggc cacagtcagc 28200
tccaggattg aacctcagaa ggtggcacag acccctcccc acagcatgca atcagaactc 28260
cgcagagctg gcgggagagt tccccctgag ggagaaatga aaacgcttgg cttctcactg 28320
ttgtcgtttt ggaggagagt gtgagagagt gaagatacag gggaggtgga aagagaaaac 28380
cgggaatgga gaggaagtga ggaagaaaga agcaaggcag acaggcacag aaagtgagac 28440

gcagagaaga aactgcgagg aagattggag atggagatgg acagagggtc cgtgaaggaa 28500

aatttcaggg agggggggagg aagagaggag tagagtgaaa aaaagaagag agagcaagag 28560

gaaatggggg aggggagaag tctcagaggt gtttagataa cacactggaa cctttgagga 28620

agagtcaggc aggagaaaat ggagctgcca aagcaccttg ttgctgtaag caaatgtgca 28680

ggaggcaagg gaagaggcca gcccaccct ctgcctctgt cctccatcca gggctgaggg 28740

ccactaggtc actgctcttg acccgctgcc attcctctga agcctgcaaa gcctactgtg 28800

tcccagcctt gctctgcatc tctgtccctt cccatgagcc aggctgggga aggagtgtct 28860

gggaggagga attctgagag ctatagagga ccccaagaga cgacctgagg gaggaaagca 28920

ggcacgggaa tggctgcaca actaactaca agggcgaggt gggatggggg caggtagaca 28980

gatgaaggga cagacagact gtcaagcagg cagaccgcac ttgaggaggg ttcccaagtt 29040

cctctgtatc atgcaccttc tgtgtctctg tccagagagt tgggggaaag ctattcctgc 29100

cctcagcctc ttcaccttct ctctaacctg aagctgtcag cttgagaata cctggtggcg 29160

ctgcctaaac cctctcaggt ctttttcccc ctggtaacaa tcattaggac tttactgttg 29220

ggctccagct tggctccagg atgaaaggaa agctggtcgg gagtgggagg tgggggcgg 29280

ggggagcagg gctgcttaca tgccctcctg cagctgggga gttagaggtt ggacagaaca 29340

cctctggccc tccccaaaca gccatggcag cccaagctcc tccctggcac acgggcaggc 29400

tctgaacccc actccagctg ctgccgactg tgcctttaaa ttactgcttt cggagggtgc 29460

atctaggggg aggtgggagg caggaaaaca tcttcctacc ggtctccctc ctccctcccc 29520

tccaagactc ttcagggttt ggagagtgat tgctgcccag agagaatctt ttccagctcc 29580

cagctggcag gctaaggccc tccggagccc aaggcgagcc caagcagaag ccagtagggt 29640

tatctgtgtc aggatcattt ccaggggaat agttctggcc cctggcaggt aaagacaggc 29700

cagaggagaa gaggcagaag aggagagaaa gcaggctctt ttgcgagcag cccaggttgg 29760

agaaaggctc tgtacttttg gcgttcctgc agggatatcc cctctcacat tggcagccag 29820

gctgagaaag ggcttcaaga tccccgcaga atgacaactc ttgttcctgc aaccctctcc 29880

ttccttcttc tctggaccct gccagggcag gtcctcctca ggtgagctga gtgggggtga 29940

gaaggaggtg catgcaaaca ctagccaagg gggatgggca acctgggcat ctggcacact 30000

ttcctctaag tttgcacaaa gcccgcagtc tggagggctt caggggagga cacgacctcc 30060

cccagttagg accccaagag ctttttatta gagatgtcaa ggcaccggtc atgctgggaa 30120

tgtttctgca caggtggaag cttcctaggc ctgggtaggg gctaggcaac taagccccag 30180

aatggagggt gtctgcccag ccacctaacc ctgtttccct atcatctggg ccacacactg 30240

ctcagatcgc taaaaatgat ggcaacatct gtcaactgcc agcccaagag gcttcaggta 30300

gcaggggttt gccccaagac cacatcagga tggagctggg ctagacattc tttggccttc 30360
tgcagtcacc ctgcctgtgc tcacttggtt gctccctatt ccttgagggg gaaaaagtag 30420
cagtcacttc ctgggcttgc aggaaagaca ggctttgggc tccctgtaaa tgttccccca 30480
ccaattgccc ctttctcatt agaccaggag ccagcctagg ttggcagggt gagctgaaca 30540
ggaaagatgg ctgtagcctg ggctgacctg atcgtaagag gagtgcctcc caaggtggag 30600
ggggtcagta ccactgcggc atactcctcc tcctcctcct gttaatccaa gaaacatcgc 30660
agggctgctt gcgtctgcaa agcctcctgc cccactgttt gcagacacat tggctggcag 30720
gcaggcaggg atgcagagca gggcaataaa cttctaaaga gtgaaaagca gctcaaagta 30780
gaatggcagc acagctggag ttgacaggag gatgcaggca gaaggggggca ccacactggg 30840
ccttgcaggc cctccaggag gggaggctgc ctcaaagctc aagctccagt agcagcattc 30900
ccactggggc ttggaaagca gaggcagtta gagcaaaacg ggcccttaga gagtctggag 30960
tccaagtcct tcgttgtaca tacggaaact gaggcctaga gatggggaat gacctgccca 31020
gggtggctca gctacctagc accagagtta gacttctctc cctgcgccac cccaagactt 31080
tacatgagaa agactgtcag ctctggatcc tcctaggtcc tgaaggtgtt taggactttc 31140
caggaatgaa agagaggaag gagtggctgc acacggtgac tctgctctca ggtcctcaga 31200
caagttagcc ccagaagagg gtactgcaac aggggctggg gaccaagact gctaagccca 31260
tagcctgtcc caggatacta ggtctaggct tagcaggaga ggtgacctga gagtggggct 31320
gaggacggcg gggagggtcc agtttctcca cgtgcaggac catgtggcct cccagcaggc 31380
aggtggcagg tggccagaag cagcagcagc acctggctag cgtggacatc ctgtggtgaa 31440
gtcacagact tctgcttgct ccagcaccac caccagcctc cttggatttt cagagcctgc 31500
ctctgtcccc agggtaaggg cttcctcttc cttagaagat tcaaggaggc aggactctag 31560
acggtcaagg tggggctgcc actgaccttc ccctactttg gctccttggc agcccaccct 31620
gggctgcctg caccatagta ttgtgataga tagagctagt tttgaagtca gtagctcttg 31680
atttgaatcc cggctttacc actcactggc tgcatgatcg tgaagaattc acctcacctt 31740
gctgatcctc agtcttctca tttatataat gggggtaatc ataatactta gttgagagga 31800
ctaatgtgag aactaaacta gacagagtgt gtaaagcatt tagcatgggg ccttagctat 31860
ttttattatt ctatcggtca ctggatagta catggatcta ctgggaaaga ggagcccact 31920
ggctagaaga agactagtcc ccagcaggac agcagttacc ctgctgccca gtcccacagt 31980
gcggcagcgg cagtgacagt ggtggagtcg gctatagagc aatgacagca ggcaggaggg 32040
agcaggaggt gcctggccca gcctggagtc cagctgagct cagccctgag cactgctcct 32100

gtccccaagg ctttgctaag ctcctagtcc ctagtttatc tgttagctgg caagttggta 32160
agtaggcata tgaaaggggc tttgctgcta ttagagggaa acccgttagg ggaaaagggg 32220
gctggcttct aagctatgcc aaagcacact ggctccttcc tcaaaatcaa gctgttagcg 32280
gggaaaatac gggagttgac ctcagcatgt ctaccctacc aggaggcttt ttcccatccc 32340
tagagtcctg cttccccacc atagttgcca aaactagttc acctgcctgc agtggcatgg 32400
tccttaacta aaacgtagct caaacatagc ttcctgttgg agcacacacc agaagccaaa 32460
actccctctt cagaagctgg gtcctacttc ctgagcagcc cagcataaag cccaagcaac 32520
tgaagccaac tctcagaaat acaagttgtc cttgagccct tcccatcctt gactaggcta 32580
gacgggcaaa actttgtagt agccacatca tttcccgcaa gacatcactc attgagtgtt 32640
ttccaagtac cttcagaaag agaacactgg acgaggcctc agcctagggt ttgagtcctg 32700
accctgtcac tagctctgag actttgaaca agtcagtcct ccttttgaag cctcagtttt 32760
ctcatctgta aaatgggaat aacagtcctc actcccatct cactgagaat tgggagggta 32820
agaagaccta ctgtctgtga aaacacttta ctagtactaa ctggcaggct gccgtgttga 32880
cttatggcgt ttttgttctc attttgcaaa tggcacccag agacccattg tcttctatcc 32940
tcctgcctga ggggtatgga aactgagagt ctgtgataag tgcagcacat aactggcctc 33000
agacttttt acctctcctg tccagggtgg ccttggcaaa agaggaagtc aaatctggaa 33060
ccaaggggtc ccagcccatg tccccctctg atttcctaga caaacttatg gggcgaacat 33120
ctggatatga tgccaggatt cggcccaatt ttaaaggtaa gaaatcttca tcctataaaa 33180
actcctccct ccccaccacc tttggcagca cactagccag ccagccccta ttgccttccc 33240
taaggaaagg aggctggagg tccagggctg gggctggttc taacagctca gaaggagcct 33300
cctcacccca ttgtccatgg gcctgtgtat ctggcactga tttctcccag cagtgccctg 33360
gcctgcagct cagggaagga gttctggaag ctgctcccag cctcctggaa tgccctgcca 33420
atggcctgtg ctttgctacc ctcaggccca cccgtgaacg tgacctgcaa catcttcatc 33480
aacagtttca gctccatcac caagaccaca atggtaaggg atctccctgc tccccacttc 33540
cagcctagtg tgagtgggag gagcccacca gataacaagc tggcatcttt tgccatatca 33600
gcccaaggaa ggttccttct ccagtgaaaa atgcctactg cccctgagat gtgttcccaa 33660
cattcctcct catggccctc ctgcccctac aggcttgctg ggcccctggg aatggcaatg 33720
tttctgaagg gcccatatct gcaccctccc aggactaccg ggtgaatgtc ttcttgcggc 33780
aacagtggaa tgacccacgc ctgtcctacc gagaatatcc tgatgactct ctggacctcg 33840
atccctccat gctggactct atctggaagc cagacctctt ctttgctaat gagaaagggg 33900
ccaacttcca tgaggtgacc acggacaaca agttactgcg catcttcaag aatgggaatg 33960

tgctgtacag catcaggtgc accgggtgga tagccaggag agtttgcgtt aaggggaaga 34020
atttggatga agactgaggg gtgggaggag ggttcccttg cactgtgatg ggccccagta 34080
agccgatgtc acttctttct tactgtcccc actaggctga ccctcatttt gtcctgcctg 34140
atggacctca agaacttccc catggacatc cagacctgca cgatgcagct tgagagctgt 34200
aagtgtctat atagagtcac agagagcctg gaataacagc cagttaaccc gacctcatga 34260
catcacagac aggaaaaggg ccactcagag ttgggaggaa cttactcaag cacacccagc 34320
caattcttgg cacagatgag actagaaacc aggggtccct aacccctagc ccagtgtgtt 34380
tttcatcccc acgttatttg ctgatctggc tgttccatgg ggctggggaa cgggaaggca 34440
tcagccttac gcacaaatgt gtttgtgctg ttttgggggg cgatttgagc gacatcttct 34500
ctgcccaaac gtatgtcccc tctcacagtg cctgaagtag ttccaccttt acctctcttg 34560
aattaagaca ttgaagatgc ctataaaaat gccagagtgg gcagggtggg ggcactcgtg 34620
gagccctgat tccccacaaa catagataac ctttgtagag aggtaggaga tgacagttct 34680
aggcacaaga cagtgcctag aatgaagagg tcagccaagg gctcctctgt ggcaggggtg 34740
ccagcctgga ggggctgccc aaagtgcaga ggaaccaagg atgacaaggt tgggtcatca 34800
cttatttcca ctccctcctc atctcctggg tcaggagaga ctgttagatc tggagccatc 34860
cactagatgt tcacagaaga aggcttataa ggctgtttcc agcataaaat aactgtgatg 34920
aatagaatgt gtcctatgtg gtcagctctt tagagttctt aaaccacttt cacatccagg 34980
ctctcaatat cccttcatgg caacgctgaa tagtggggag agcagctact ataattccca 35040
ttttacagaa ggggaagcga cttacccaaa gtcacagagc tactctgggt caaggcccaa 35100
cactggatgt ctgtgagcct tggtgtttct gctcccatct ctacctgaca tgtgcaaata 35160
tggtcatata cgtggcatga gccctcatga actgctatgt atgtgtgaat aacttgacat 35220
gttcagacat tagctttaaa tttgccccaa ctcaaggccc tggctaggat ctatggtatg 35280
cacattgtcc catgcatgac tggcaggata taaatagctg tctcctccct gaggtgatgt 35340
ggcagaagtc ctggctctga ggctagctca gcttaagaca tgtgttcggt acacatgctt 35400
actcccctcg gcttgagata tggcaggagg ctcccagctg tgttccttga catctttcca 35460
accctcttga cctgctttct aggtctgtca gaataatttc tctttgggct tttctcaagc 35520
ttgagaacag ggtaggagag gagtctgtaa tcttcctgat ttacctgccc actcaggatg 35580
ggcccatgag acaactgacc aagtgtcctc tccagaaaag ttttcctttc cataatgagc 35640
atgaagctcc taaaccccag accttactcc tcacctgttc tctcctgcct gttctgatcc 35700
taccctacc cccacccagc atccatactc tgcagccctc tgccatctct gtcactttca 35760

gttggctaca ccatgaaaga cctcgtgttt gagtggctgg aagatgctcc tgctgtccaa 35820
gtggctgagg ggctgactct gccccagttt atcttgcggg atgagaagga tctaggctgt 35880
tgtaccaagc actacaacac aggtaaaaac caggatcttt attggctgtg agccacagga 35940
gaaccttcag atgtgaggca aggggacaag gcagctggcc tctatcgaag gttcaatcta 36000
tgtcagatat ggggccaggt gcttcatata acttatttat ttgtgacagc aactataaat 36060
accatttttc agaaagggaa aactgaggtt ctaagaagtt agttaaatca agtgtccaag 36120
tcgcacagct agtaaatggc agatttagaa tttaaatcca agccttctaa ctccaaagtc 36180
tgtgttcttg ttcctgtcgc atgctgcaga gcaaatgaca gtgggtggac agtttcccac 36240
ttttcctggg gtcaactttg tgtggaagcg gggaatccac atctggtaat agggagtgac 36300
tatgctgaac ccagagggaa caagcctgag ggcagtggcc agcttacagt ccacacctcc 36360
cactgtcgct accactgcca gctctcagga acttcccagg tgtgaaatgt cagctaagaa 36420
gtggagaagg gctgctttcc caggctgagc atctgccccc ctgcagcttc tgtaagagtt 36480
gctggctcca cattttccct gacaatttgt ctgcgcggag ctctgaccct gaaatagagg 36540
cagcagcctg aagaggggag atcccgggtc ccttccatgg tgggctgata tgttttgggt 36600
gctgccccat gaatatgtgt gtgcatgtat gtgggggtgt ggggagtggg aagaggaaaa 36660
ggccattgaa tttagaaaat gcaaggggga atgaaaattg atataaactt tttggagtaa 36720
aaaagttggc aatttgtata tagatcagtg gttttcaaac tttggtgtgc actatattcc 36780
cctggggagt ctattaaaat gttgcctcct gggctctacc tctcagagat tctgattatg 36840
gttacctgta cttggagtga tgcctagaaa tctgcatttt aacaagcatc ctgaggattg 36900
tggtgcaggt ggtccagcaa acatgctttg agaaacttta gtctagaagc taaaattgtg 36960
tttctactta tttatccagg aatcctactc cttgaaatgt attatatgga aatattccaa 37020
aatactgcat tggccttaga ccccaagggg gctcctgggt ggctgtttat attagcaaaa 37080
acttggaact gttttaaata tccaccaata gggcagaggt taagaaaatt acatgtgcgt 37140
ctgcttggtg gactagtatg cagacattaa aaggaatggt tacaaggtgc ttatatcaat 37200
gtgggaaagt aggagtgaaa aaaggtaaga ttcggaggta taaatacaga atgattatat 37260
tatggtaaaa gaaacaaacc aaaagcatgc atttaaaaac cttaggtgga aatagcacta 37320
aatgtctcta aagtaatgtg atacaacaga taagagcaag aactttggag gcagatagag 37380
ctgagtttgg attctggctc catcatttac tagttgagcg cccttgagaa atataattta 37440
atctcttcat gcctccgtat tcctatctgt aatatgagag ataataattt tacttgcccc 37500
aaagagttgc tataggatta aatgaggtaa agtcaggtcc cttagtaagg gtttgataaa 37560
tgtaagctta tgaatattat tattccttct ctacttctta ttttttctat actaaacaag 37620

ttttaatttt ggaattaaaa acaaaactgc tattgcaaaa aaaagaagaa gaagaaggaa 37680
agaaaggaaa gaaagtatag gtacctaata aaggcatgtc acggcaagaa tcccagggtg 37740
ggatgaagtg aggtggattc agggaacagt aagaggtcac agcatcatac aatgtgtcaa 37800
agccataatc agtcattgac aacatcactg actttgaaga aacaagtagg cctgataaga 37860
ctagacctag gccagaataa aataacaata tatgcatctt ctccttcttg cttggattat 37920
cctgagcctc tgcttccaag agaagctctg cgttagtatt ttcttagccc tagggccctg 37980
gaaagcagag taaagtgaag tctggcaaag gaggtggagc attggtcggt gtgcctgggc 38040
caatgcactt tttgttggta gaatagaatg ttgcagcctg atgctctatg ggaacctgct 38100
ccccgctgtc tggagctggt ctctgttggg cactgttttc agatgtttgc tatccggctc 38160
ttctcacttg cttaagacac aagcagtgtt tatcccagtg aattgttgtt gctagagagc 38220
ttctggggca ctagcaagca caagcatggc ccctgagatc taagtggagt ctaagctaat 38280
tgttttgtta ataatgataa caatagtaac agctacctat gtgtcaaaca cctcaaatac 38340
attatcttgt ttaatttttc caatagccca gtgatgcagt tggtatcact actctgtttt 38400
acaggcaagg aaactgaagc tcaaaacgct tgggtgtttt gcccaaagtc acacagtgga 38460
aggttcagga ttcaaaccca ggcagtctga ctctggagcc ctcctgctcc ttatcctctt 38520
gctgtgcagt cacgtgaaag gccttccgct ggaaatgaaa gtgctgggag ccaaaggcca 38580
agtctttgat ggtcgctctg agagcttcct tagtacgtac atctggcatt acctggggcc 38640
attatgaaag gctcattcat gggcccagct gaggtgccca gaggctggtc agtagagggc 38700
tgccctaatc ctgtcctaat cccctcaggc tcagctacag gctgggtacc tggggcctga 38760
cacctcacct ctctccttgc agggaaattc acctgcatcg aggtaaagtt tcacctggaa 38820
cggcagatgg gctactatct gattcagatg tacatcccca gcctactcat cgtcatcctg 38880
tcctgggtct ccttctggat caacatggat gctgcccctg cccgtgtggg cctgggcatc 38940
accaccgtgc tcaccatgac cacccagagc tctggctccc gggcctcttt gcctaaggtg 39000
aagagacatg caagggaact tgcttgccat agagtacccc atttcccatt catcccaatg 39060
tgccctcctg ctcaccattt accttggatt taaggttgag aagacccact gggcaaaagc 39120
tggtgtcctg cattctctct ccctgtaata tactgctaat aacattgttg acaatgtgga 39180
ggctggggtt agggatacta tctggctagg gaatgaggag agacagtgcc ttactttgac 39240
tctcttcagg aaatactagc atgtccaagg tcaaaagatc ctcagcaagt agaaagtcct 39300
actgcctcat aatacaaatg aggagaatac agcccagagt ggtaaaataa ccttcttaag 39360
atcatgcctt ataagaaaca accaaaacaa tatcagagtt ttgaaccaat ttctgtcatc 39420

ttactgtatc agaaattaga aactatcaca aagcgctcta gagatattat cctttctttt 39480
gggatgagaa aattgaggta gaaaatagca attaaagtca gtgaactcca gttgtctggg 39540
cccagtgatg ttttgattac gcaggaagat ttgtgctgat gggactgggc acagattctc 39600
cctgcctcat gtcttcatgt agggaattta gaggtcacct caaagtcctt acagtcagca 39660
aggctctgaa cgaatcaggg aacaaggcca tttattcatt caataaatat ttattgagca 39720
tgtactatgt accagccact gaactttgtc atggaaaaat aacaggaagc aaaagcaaca 39780
cagttccttg tggatc 39796

INTERNATIONAL SEARCH REPORT

Inter nal Application No
PCT/EP 00/02600

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C12N15/12 C07K14/705 C12Q1/68 G01N33/68

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 C07K C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)
CHEM ABS Data, BIOSIS, SCISEARCH, MEDLINE, EMBASE, EPO-Internal, BIOTECHNOLOGY ABS, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	MATZENBACH BERND ET AL: "Structural Analysis of Mouse Glycine Receptor alpha Subunit Genes: Identification and chromosomal localization of a novel variant, alpha-4." JOURNAL OF BIOLOGICAL CHEMISTRY, vol. 269, no. 4, 1994, pages 2607-2612, XP002142164 ISSN: 0021-9258 abstract; figure 3 --- -/--	1,7-9, 11,13,16

[X] Further documents are listed in the continuation of box C.

[ ] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
10 July 2000	24/07/2000

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL – 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Gurdjian, D

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

Inter ›nal Application No
PCT/EP 00/02600

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
A	GRENNINGLOH G ET AL: "ALPHA SUBUNIT VARIANTS OF THE HUMAN GLYCINE RECEPTOR PRIMARY STRUCTURES FUNCTIONAL EXPRESSION AND CHROMOSOMAL LOCALIZATION OF THE CORRESPONDING GENES" EMBO (EUROPEAN MOLECULAR BIOLOGY ORGANIZATION) JOURNAL, vol. 9, no. 3, 1990, pages 771-776, XP002142165 ISSN: 0261-4189 abstract; figure 2 ---	1,7-9, 11,13,16
A	HILLIER L. ET AL.: " zu62b06.s1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:742547 3' similar to TR:G817957 G817957 GLYCINE RECEPTOR SUBUNIT ALPHA 4" EMBL DATABASE ; ACCESION NUMBER AA400068, 23 April 1997 (1997-04-23), XP002142166 the whole document ---	1,7-9, 11,13,16
A	DOYLE JL (REPRINT) ET AL: "Ataxia, arrhythmia and ion-channel gene defects" TRENDS IN GENETICS, 1998, 14, 92-98, XP002142167 abstract; table 2 ---	1,7-9, 11,13,16
A	BECKER CM ET AL: "ISOFORM-SELECTIVE DEFICIT OF GLYCINE RECEPTORS IN THE MOUSE MUTANT SPASTIC" NEURON, 1992, 8, 283-289, XP002142168 abstract -----	1,7-9, 11,13,16

1

Form PCT/ISA/210 (continuation of second sheet) (July 1992)